INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



GENES TRANSPONIBLES EN EL MAIZ

Agosto 1984
350 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 95
Agosto 1984

INVESTIGACION Y
CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

EDITA Prensa Científica, S. A.
Calabria, 235-239
08029-Barcelona (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Fernando José Rodilla: *Armas antisatélites*; Elías Zamora Acosta: *Metalurgia superficial precolombina*; Salvador Reguant: *Extinciones masivas en los océanos*; Enrique Sánchez Monge: *Elementos genéticos transponibles del maíz*; José Bernabeu: *Preferencia atómica entre izquierda y derecha*; Manuel Puigcerver: *Estructura y evolución de los granos interestelares*; Julio Muñoz: *Tuercas y pernos*; Luis Bou: *Juegos de ordenador*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

Ismael Fernández de la Cuesta

Libros:

Pedro Laín Entralgo y Guillermo Olagüe de Ros



LA PORTADA

El dibujo de portada muestra la notable variegación que puede darse en la pigmentación de las semillas del maíz por acción de los elementos genéticos transponibles, fragmentos móviles de ADN descubiertos por Barbara McClintock a finales de los años cuarenta. Algunos de esos elementos se han aislado y caracterizado hasta el nivel molecular (véase "Elementos genéticos transponibles del maíz", por Nina V. Fedoroff, página 44). Cuando los elementos se desplazan de una posición a otra del genoma producen mutaciones y reestructuraciones cromosómicas, con lo que afectan a la expresión de los genes. Por ejemplo, la inserción de un elemento en un gen que regula la síntesis de un pigmento lo inactiva y la semilla pierde el color; la posterior escisión del elemento de ese locus restaura la función: en la semilla aparecen manchas pigmentadas. Las semillas representadas en el dibujo son una muestra de distintas mazorcas.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
08029-Barcelona (España)
Teléfono 322 05 51 ext. 33-37

Condiciones de suscripción:
España:
Un año (12 números): 3850 pesetas
Extranjero:
Un año (12 números): 33 U.S. $
Ejemplar atrasado ordinario:
350 pesetas
Ejemplar atrasado extraordinario:
450 pesetas

Distribución para España:
Distribuciones de Enlace, S. A.
Ausias March, 49 - 08010-Barcelona

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - 08029-Barcelona

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - 28030-Madrid
Teléfono 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.º - 08034-Barcelona
Teléfono 204 45 83

Controlado por O.J.D.
OJD DIFUSION PUBLICIDAD

PROCEDENCIA DE LAS ILUSTRACIONES

Pintura de la portada de Enid Kotschnig

Página	Fuente
8	Fuerzas Aéreas de los Estados Unidos
9-18	George V. Kelvin
21	A. I. Goldberg (*arriba*), Robert Sonin (*centro*), Metropolitan Museum of Art (*abajo*)
22	Metropolitan Museum of Art (*arriba*); Museo del Oro, Bogotá (*abajo*)
23	Alan D. Iselin
24-25	Heather Lechtman, Instituto de Tecnología de Massachusetts
26-28	Alan D. Iselin
31	Steven M. Stanley, Universidad Johns Hopkins
32-39	Patricia J. Wynne
44-46	Fritz Goro
47-50	Bunji Tagawa
51	Nina V. Fedoroff, Institución Carnegie de Washington, y Bunji Tagawa
52-54	Bunji Tagawa
57-65	Walken Graphics
68-78	A. Escarré, C. Gracia, F. Rodà, J. Terradas y Ricardo Génova
81	Observatorio Nacional Kitt Peak
82	John Moss, Black Star
83-86	Allen Beechel
87	John Moss, Black Star
88-91	Allen Beechel
93-97	Walken Graphics
98	National Machinery Co.
100-104	Alan D. Iselin
105-109	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - 08005-Barcelona
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S.A.
Nápoles, 249 - Tel. 207 40 11
08013-Barcelona
Printed in Spain - Impreso en España

Los autores

RICHARD L. GARWIN, KURT GOTTFRIED y DONALD L. HAFNER ("Armas antisatélites") son, respectivamente, becario del Centro de Investigación Thomas J. Watson de la International Business Machines Corporation (IBM), profesor de física de la Universidad de Cornell y profesor adjunto de ciencia política del Boston College. Garwin se licenció en ciencias por el Instituto Case de Tecnología y se doctoró en física por la Universidad de Chicago; permaneció tres años en esa institución, tras lo cual se incorporó a IBM, donde ha desempeñado diversos cargos, entre otros el de director del Laboratorio Watson. Gottfried nació en Austria y emigró a Canadá en 1939. Se licenció por la Universidad McGill y se doctoró por el Instituto de Tecnología de Massachusetts (MIT). Trabajó en la Universidad de Harvard desde 1960 hasta 1964, año en que pasó a la de Cornell, donde ha permanecido hasta hoy. Hafner cursó el primer ciclo en el Kalamazoo College, Michigan, doctorándose en ciencias políticas por la Universidad de Chicago, en 1972. Durante 1977 y 1978 abandonó la docencia para trabajar en la Agencia para el Control de los Armamentos y el Desarme.

HEATHER LECHTMAN ("Metalurgia superficial precolombina") enseña arqueología y tecnología antigua en el Instituto de Tecnología de Massachusetts. Colabora en el programa de antropología y arqueología de los departamentos de humanidades y de ciencias de los materiales e ingeniería. Tras su licenciatura por el Vassar College, comenzó a trabajar en el departamento de biofísica del Instituto Sloan-Kattering para la Investigación del Cáncer. Cuatro años después se trasladó al Instituto Americano de Física. Posteriormente ingresó en la Universidad de Nueva York para continuar sus estudios, obteniendo la licenciatura en Bellas Artes y un diploma de conservadora de obras de arte y materiales arqueológicos. Pasó luego al MIT. Dirige el Centro Interinstitucional para la Investigación de Materiales Arqueológicos y Etnológicos.

STEVEN M. STANLEY ("Extinciones masivas en los océanos") resume así su biografía académica: "Soy profesor de paleobiología en el departamento de ciencias de la tierra y de los planetas de la Universidad Johns Hopkins. Cursé la licenciatura en la Universidad de Princeton y me doctoré por la de Yale. Estoy interesado en el área de confluencia de la geología con la biología. Mis investigaciones anteriores se han centrado en la morfología funcional y la evolución de los organismos marinos y en la naturaleza de la evolución a gran escala. Ultimamente me he ocupado de las faunas de moluscos de Florida que se extinguieron a principios de la última edad del hielo".

NINA V. FEDOROFF ("Elementos genéticos transponibles del maíz") se halla adscrita al departamento de embriología de la Institución Carnegie de Washington. Da también clases de biología en la Universidad Johns Hopkins. "Mi interés por los elementos genéticos transponibles del maíz arranca de un encuentro con Barbara McClintock durante una visita a Cold Spring Harbor, hacia 1977, encuentro que me indujo a leer todos los trabajos pioneros de esa autora. Precisamente cuando el aislamiento de los elementos del maíz aparecía ante mí como una idea gloriosa y poco realista se me ofreció un puesto en la Institución Carnegie. La política que se sigue en ese centro, dedicar simultáneamente atención a la investigación y apoyo a objetivos a largo plazo, permitió consagrarme a la biología molecular del maíz, que estaba entonces en sus balbuceos."

MARIE-ANNE BOUCHIAT y LIONEL POTTIER ("Preferencia atómica entre izquierda y derecha") pertenecen al Centro Nacional Francés de Investigación Científica (CNRS). Colaboran en el tema sobre el que versa su artículo desde 1974. Bouchiat trabaja en el laboratorio de física de la Escuela Normal Superior (ENS) de París. Tras licenciarse por la Escuela Normal Superior de Sèvres se trasladó a los Estados Unidos, donde disfrutó de una beca de dos años en la Universidad de Princeton. Volvió al ENS, doctorándose en 1964. Pottier comenzó su trabajo experimental en el laboratorio de Bouchiat; volvió a doctorarse en 1972.

ANTONI ESCARRE, CARLOS GRACIA, FERRAN RODA y JAUME TERRADAS ("Ecología del bosque esclerófilo mediterráneo") son biólogos especializados en ecología terrestre. Escarré es profesor titular y dirige el departamento de biología de la Universidad de Alicante. Sus líneas de trabajo se centran en la biogeoquímica de cuencas forestales y en la ecología de las comunidades de artrópodos y de vertebrados terrestres. Gracia es profesor del departamento de ecología de la Universidad de Barcelona y ha estudiado la organización vertical de los bosques en relación con el régimen de radiación en la bóveda forestal. Rodà, profesor del departamento de ecología de la Universidad Autónoma de Barcelona, ha abordado el papel de las aguas de lluvia y de avenamiento en la circulación de nutrientes en los bosques. Terradas es catedrático de ecología y director del mismo departamento, y se ha especializado en la ecofisiología de los vegetales, en ecología forestal y en las relaciones entre ecología y educación ambiental. Todos ellos participan en un programa común dirigido a esclarecer aspectos básicos de la estructura y el funcionamiento de los bosques mediterráneos.

J. MAYO GREENBERG ("Estructura y evolución de los granos interestelares") enseña astrofísica de laboratorio en la Universidad de Leiden. Durante la segunda guerra mundial realizó trabajos de hidrodinámica teórica para la Comisión Asesora Nacional de Aeronáutica. Se doctoró en física en 1948. En 1952 ingresó en el cuerpo docente del Instituto Politécnico Rensselaer. A finales de los años cincuenta, su interés se trasladó de la física a la astronomía. En 1968 y 1969 la Universidad de Leiden le ofreció inaugurar la nueva cátedra de astrofísica de laboratorio. De nuevo en los Estados Unidos ocupó la cátedra de astronomía de la Universidad estatal de Nueva York en Albany. En 1975, Greenberg aceptó un empleo permanente en Leiden, donde creó un laboratorio para estudiar los efectos de la radiación ultravioleta sobre los granos interestelares.

FREDERICK E. GRAVES ("Tuercas y pernos") es ingeniero y ha trabajado como consejero para fabricantes de elementos de unión. Después de licenciarse en ciencias por la Universidad de Pennsylvania, en 1940, se incorporó al Battelle Memorial Institute para investigar en ingeniería. Estuvo tres años en dicho instituto, integrándose luego en la industria privada. En 1969 se estableció por su cuenta.

Hace...

José M.ª López Piñero

... doscientos años

Sebastián Guerrero Herreros y Morales y Reina estudió las epidemias de viruelas que padeció la ciudad de Sevilla tras las inundaciones del Guadalquivir de 1784 y 1785. Poco más tarde, expondría el resultado de este trabajo en la última de sus publicaciones, titulada *Memorias physicas y médicas sobre las viruelas* (1786).

Nacido en la capital andaluza hacia 1720. Guerrero fue una de las figuras científicas centrales de la Regia Sociedad de Medicina, de Sevilla, durante la segunda mitad del siglo XVIII. Según Chinchilla, cursó medicina en la Universidad de Granada y, tras graduarse, volvió a su ciudad natal, en la que residió hasta su muerte consagrado al ejercicio profesional. Los únicos datos seguros acerca de su biografía hasta ahora disponibles son las noticias recogidas por Hermosilla en el archivo de la citada institución y las que el propio Guerrero incluye en sus obras. Sabemos así que en 1744 solicitó su ingreso en la Regia Sociedad, superando con éxito un examen consistente en exponer la osteología de la cadera y las causas, síntomas y tratamiento de los aneurismas.

Tres años más tarde, organizó una "academia" sobre "materias de física moderna, química, medicina y otras ciencias" que debía funcionar seis horas semanales a lo largo de todo un curso. En 1768, cuando la Regia Sociedad perdió la propiedad de los locales que ocupaba, las juntas se celebraron en el domicilio particular de Herreros, donde también se trasladó la biblioteca de la institución y diversas máquinas, entre otras, la de electrizar. Allí estuvieron hasta que, en 1771, la Regia Sociedad se instaló en el Colegio de los Ingleses, desalojado por los jesuitas tras la expulsión de éstos. El éxito profesional o la fortuna personal de Herreros tenían que ser notables porque, además, ofreció en 1770, sufragar los gastos que implicaban las llamadas "conclusiones", que eran sesiones públicas anuales en las que se defendían tesis, especialmente de carácter práctico. Con anterioridad a 1773, una de las tareas que desempeñó fue la de director en las "demostraciones anatómicas" que tenían lugar en el anfiteatro de la Regia Sociedad. En nombre de esta última, actuó también en varias ocasiones como examinador de médicos, cirujanos y boticarios.

Por último, entre 1762 y 1790, presentó a la Sociedad al menos nueve "disertaciones" de las cuales seis aparecieron publicadas en sus *Memorias Académicas*. Quedaron manuscritas las que dedicó a la histeria (1762), cuyo origen uterino negaba, a una "epidemia catarral" padecida en Sevilla (1768) y a estudios médicolegales en torno a las muertes violentas (1763) y la esterilidad (1790). Se imprimieron las relativas al tratamiento de enfermedades respiratorias con preparados de quermes mineral (1766), a la putrefacción de los humores (1772), al texto hipocrático *De glandulis* (1786) y a la terapéutica mercurial de la sífilis (1786).

De todas estas "disertaciones", sobresale la titulada *De la putrefacción de los humores, y medios de corregirla*, que fue publicada en 1772, como acabamos de decir. Partiendo de las obras de Hermann Boerhaave y de John Pringle, Herreros había llegado al convencimiento de que la "putrefacción" es un mecanismo patogénico de gran importancia. En consecuencia, se dedicó a investigarlo, reuniendo observaciones clínicas a lo largo de ocho años y realizando experimentos, en los que aseguró haberse gastado "más de veinte mil reales". Expuso sus hallazgos en esta disertación aunque, desde una actitud muy exigente, los consideró insuficientes. Intentó, sin embargo, distinguir la fermentación de la putrefacción, considerando esta última como el resultado final "del movimiento intestino de las partes vivientes, especialmente líquidas, por el cual las más sólidas se reducen a una masa pultácea y los humores a un licuamen pultáceo, de lo que resulta que las sales nativas propias del viviente se convierten en álcali volátil y la parte pinguedinosa, en una masa fétida, negra, acre, tenaz o muy disoluble". Muy notable es que considere que la putrefacción es un proceso biológico general, "uniforme en todos los vivientes del reino animal y vegetal".

La más importante de las publicaciones de Guerrero fue una obra en dos volúmenes titulada *Medicina Universal* (1774-1777). El primero de ellos está fundamentalmente dedicado a cuestiones estequiológicas y constituye, en opinión de María Luz Terrada, una de las aportaciones más rigurosas que se hicieron a este tema en la España del siglo XVIII. Estudia la fibra como unidad elemental de la materia viva, con gran cantidad de datos microscópicos y químicos y sobre la base de los criterios conceptuales de Hermann Boerhaave y Albrecht Haller:

"En todos los animales, registrados sus primeros elementos con el microscopio, se conoce que son fibras, o concrementos inorgánicos, rudos, sin formación, ni figura; y de estos concrementos se componen las fibras primeras y simplicísimas, de que se compone el cuerpo viviente sensible; y se llaman partes sólidas, firmes y consistentes. Verdad es, que en la fibra se hallan diversos elementos, pero también es cierto, que la sola fibra es materia común y verdadero fundamento del cuerpo; de suerte, que en el cuerpo viviente no haya parte que no se forme y componga de fibras, aún aquellas donde la vista no las registre, como son el cerebro, médula espinal y médula de los huesos... de ellas se componen el tejido celular, los vasos todos, las membranas, cartílagos, huesos, ligamentos, tendones, músculos, nervios, los parénquimas, carnes parenquimatosas o entrañas, los pelos y las uñas...

"Los elementos, o partes primeras, de que se componen las fibras, son unos sólidos y otros fluidos; pero unos y otros tan estrechamente unidos... que solamente los puede separar el fuego, o una maceración prolongada, y putrefacción muy larga... las partes sólidas... estuvieron contenidas en líquidos o fluidos, y de ellos se separaron yuxtaponiéndose a sí mismos mutuamente por la fuerza, eficacia, operación y ejercicio de la vida, y uniéndose unas a otras con el beneficio de un gluten tenuísimo, acuoso, o pingüe, que de una y otra naturaleza lo hay, hacen y forman las fibras mínimas... esta fibra mínima, que es invisible, tangible, visible, que es la primera parte conocida del cuerpo..."

Habiendo observado que los cadáveres quedan reducidos a polvo o tierra, y apoyándose en las observaciones hechas por Malpighi acerca del desarrollo del embrión de pollo, Herreros deduce el carácter *terrestre* de estas "fibras mínimas", aunque con características particulares debidas a su condición de partes vivas. Esta parte térrea la considera unida por un gluten, que actuaría de elemento fluido de su estructura, en cuya composición entrarían como elementos, además de la "tierra", "agua, aceite, hierro y aire".

Se adhiere al concepto de "tejido celular" propuesto por Haller como estructura fundamental del organismo:

"Llamaremos tejido celular, a aquella amplísima porción del cuerpo humano, que se compone de los dos géneros de fibras descritas ya que son los estambres y las láminas, bien mezcladas entre sí en diversas proporciones, o ya hecho al tejido de solas fibras, o ya de solas láminas..."

Expone ampliamente esta doctrina textural, de transición entre la teoría fibrilar y la noción de tejido de Bichat. Por último, la utiliza como fundamento de la fisiología y, sobre todo, de una concepción patológica general explícitamente morfológica, según la cual la inflamación, los trastornos circulatorios locales, los tumores, etc., serían primariamente afecciones de dicho "tejido celular":

"Considero yo el tejido celular como universal teatro de las funciones animales, vitales y naturales en todos estados: para red es esa tela, o tejido de células o cedillas por donde todo pasa y donde tantas cosas se reciben.

"La experiencia de los cadáveres ha demostrado que siempre que hay inflamación la hay en el tejido celular, y no siempre en los vasos serosos; no obstante que muchas veces hay inflamación en los vasos que se junta con la inflamación del tejido".

Tal como Riera ha puesto de relieve, desarrolla con especial detalle la aplicación de este enfoque teórico a los tumores.

El segundo volumen de la *Medicina Universal* está íntegramente consagrado a la viruela, enfermedad que ya había estudiado en el primero desde un enfoque general. En la década siguiente, como hemos dicho al principio, Guerrero Herreros volvió a ocuparse del tema en sus *Memorias physicas y médicas sobre las viruelas* (1786), libro que ha sido generalmente ignorado.



481

JUEVES 17.

DISERTACION PHARMACEUTICA

DE LOS VARIOS EFECTOS
que producen los preparados mercuriales hechos con Sales, ó Azufres, explicando el mechanismo, y causas de esta diferencia.

POR

DON SEBASTIAN HERREROS
Socio Medico de Numero.

1. NO menos de tres siglos que tuvo en Europa casi sin saber como su principio aquel azote de Dios, que como otro cruel Atila, ha desfigurado, despedazado, y muerto mas indivi duos de nuestra especie, que quantos tormentos, y maquinas belicas inventó la inter-

1. Grabado alegórico utilizado en las **Memorias Académicas** *de la Regia Sociedad de Medicina y demás Ciencias, de Sevilla, en las que publicó sus trabajos científicos Sebastián Guerrero Herreros. 2. Página del trabajo de Herreros sobre la terapéutica mercurial de la sífilis (1786).*

Armas antisatélites

Si no se toman medidas para contener el desarrollo de esas armas, se verá seriamente amenazada la aportación positiva de los satélites a la seguridad internacional

Richard L. Garwin, Kurt Gottfried y Donald L. Hafner

Los formidables avances de la tecnología militar se han traducido en una fuerte disminución de la seguridad mundial. Hace cuatro décadas se habría necesitado varios meses para doblegar a una gran potencia por medio de un ataque total. Hoy podría borrarse del mapa toda una civilización en cuestión de horas. Los satélites artificiales son un raro ejemplo de nueva tecnología militar que, en términos generales, ha mejorado la seguridad mundial. Sin embargo, de no adoptarse rápidamente medidas políticas, el desarrollo del arsenal de armas antisatélites (ASAT) no tardará en conducir la carrera armamentista a una era en la que los satélites perderán la condición de relativa seguridad de que hoy gozan.

La amenaza que se cierne sobre los satélites tiene su origen en el doble papel que éstos desempeñan. En tiempos de paz, proporcionan tanto a los Estados Unidos como a la Unión Soviética una información rápida e irremplazable de las fuerzas y operaciones militares de la otra parte. Consecuencia de ello es que son esenciales en muchos aspectos de la planificación militar e indispensables para la comprobación del cumplimiento de los acuerdos sobre control de armas estratégicas. En la hipótesis de una crisis política en la que se viesen envueltas ambas superpotencias, conseguir que no se produjese un conflicto armado dependería en gran medida de los satélites, como instrumentos necesarios para la evaluación de las acciones del otro bando. Si se produjese una ruptura de las hostilidades, todas las posibilidades de evitar una escalada catastrófica dependerían del flujo continuado de la información de vigilancia y de las comunicaciones procedentes de los satélites. Pero la extraordinaria capacidad que éstos tienen para ver, escuchar y comunicar aumentaría enormemente la eficacia de las fuerzas militares en tiempos de guerra. Por esa razón, los satélites se convertirían en objetivos especialmente tentadores tan pronto como las hostilidades pareciesen inminentes. He ahí la razón del enorme interés que ambas partes han venido mostrando por el desarrollo de armas antisatélites.

En los años sesenta, Estados Unidos y la Unión Soviética desarrollaron armas con capacidad antisatélite. Estados Unidos acabó desmantelando su propio sistema antisatélite y adoptó la postura de que lo más beneficioso para su seguridad sería abstenerse en adelante de todo tipo de competición en armamentos de esa clase. No así la Unión Soviética, que continuó realizando pruebas esporádicas de su propio sistema.

A mediados de la década de los setenta, Estados Unidos cambió de postura, aunque manteniendo como objetivo último lograr un medio espacial libre de armamentos. Esta nueva política tenía dos componentes interrelacionados: del lado militar implicaba necesariamente desarrollar una nueva arma antisatélite más compleja; del lado de la diplomacia, implicaba también la necesidad de negociaciones dirigidas a la firma de un tratado que acabara por prohibir todas esas armas. En 1974 comenzaron a celebrarse conversaciones bilaterales sobre este tema, pero Estados Unidos las suspendió en 1979 a raíz de la intervención soviética en Afganistán. Desde entonces, la Unión Soviética ha sometido a las Naciones Unidas, en 1981 y 1983, sendos proyectos de tratados pidiendo el control de armamentos en el espacio, sin que Estados Unidos haya dado respuesta a tales iniciativas.

Ambos países han proseguido entretanto sus trabajos en el campo de las armas antisatélites. Más aún, la competencia en armamentos con base en el espacio seguramente se incrementará si Estados Unidos sigue la recomendación del presidente Reagan y se embarca en una exploración a gran escala de las posibilidades de un sistema de misiles antibalísticos con base en el espacio, denominado oficialmente "Iniciativa de Defensa Estratégica" y conocido comúnmente como el programa "Star Wars" ("Guerra de las estrellas"). En realidad, dada la íntima relación tecnológica que existe entre armas antisatélites y defensa por misiles balísticos, no podrían cumplirse ni siquiera las fases preliminares de semejante programa, si el tratado en vigor sobre misiles antibalísticos se completara con un acuerdo similar sobre limitación de armas antisatélites.

La entrada en escena de armas contra satélites pondría en evidencia que el papel positivo que los satélites desempeñan en el control de armamentos, en la creación de un clima de confianza y en la solución de conflictos se ha venido considerando de menor importancia que su capacidad de apoyo a auténticas operaciones militares. En lo que sigue trataremos precisamente esa cuestión. Además, no hay que dejar de señalar que queda todavía una oportunidad de negociar una limitación, militarmente importante y posible de comprobar, al progreso de la tecnología antisatélite, lo cual redundaría en interés de Estados Unidos y del mundo entero.

Hay que analizar muchos problemas antes de emitir un juicio sobre si es deseable o factible limitar, por vía negociadora, el progreso de la tecnología antisatélite. Es necesario evaluar los papeles de los satélites militares existentes y su vulnerabilidad a las armas antisatélites, así como el impacto probable de dichas armas en diferentes tipos de crisis. En este artículo abordaremos todas estas cuestiones y otras con ellas relacionadas.

Conviene distinguir entre los diversos tipos de satélites militares; y ver también en qué modo dependen de ellos Estados Unidos y la Unión Soviética. Las órbitas de la mayoría de estos

satélites se encuadran en cuatro categorías: (1) órbitas bajas, prácticamente circulares, con un período de unos 100 minutos y una altitud que va desde un centenar hasta varios miles de kilómetros; (2) órbitas geosincrónicas, en las que el satélite gira alrededor de la tierra al compás de la rotación de la misma y permanece fijo sobre un punto determinado del ecuador, a unos 36.000 kilómetros de altura; (3) órbitas altamente elípticas, que descienden hasta escasos centenares de kilómetros en su cota más baja sobre el hemisferio sur y se elevan hasta alcanzar su cénit, a unos 40.000 kilómetros, sobre el hemisferio norte, y (4) órbitas sincrónicas, que son prácticamente circulares en un plano inclinado y con una altitud de alrededor de 20.000 kilómetros [*véase la figura 3*].

La altitud del satélite tiene que ver con la misión que le haya sido encomendada. Desde una órbita baja el satélite alcanza la visión más detallada de la superficie terrestre, pudiendo también llegar a detectar débiles señales electrónicas procedentes de la tierra, del mar o del aire. En consecuencia, los satélites de fotorreconocimiento y de vigilancia oceánica, y la mayoría de los de inteligencia electrónica, se sitúan en órbitas bajas. Si se les encomienda vigilar extensas zonas o comunicar con una estación fija en tierra, es preferible una órbita sincrónica. Los satélites norteamericanos con sensores diseñados para proporcionar una alerta precoz de lanzamiento de un misil soviético ocupan tales órbitas, así como la práctica totalidad de los satélites de comunicaciones de Estados Unidos.

Dado que la Unión Soviética cuenta con importantes instalaciones en la región ártica, donde es difícil alcanzar una clara línea de visión para un satélite sincrónico situado sobre el ecuador, los rusos han recurrido a órbitas altamente elípticas para muchos de sus satélites de comunicaciones de la clase Molniya y de alerta precoz. Los satélites del sistema norteamericano de datos sobre satélites (Satellite Data System, SDS), que tienen encomendadas las comunicaciones con fuerzas desplegadas en el Artico, están situados en órbitas semejantes. Un satélite instalado en una órbita de ese tipo, y cuyo cénit se halle a grandes altitudes sobre el hemisferio septentrional, permanece visible para su estación terrestre por espacio de ocho o más horas de su período de doce.

Otra importante misión militar a la que contribuyen los satélites es la de lograr una navegación global más precisa. Los primeros satélites de navegación global ocupaban órbitas bajas, pero el sistema norteamericano de posición global (Global Positioning System, GPS), más perfeccionado, y otros sistemas soviéticos semejantes se están desplegando ahora a unos 20.000 kilómetros, altitud en la que basta con un número relativamente reducido de satélites (18 o 24) para tener simultáneamente cuatro de ellos al alcance de la vista desde cualquier punto de la tierra. El sistema integrado de detección nuclear a bordo (Integrated Onboard Nuclear Detection System, IONDS), para detectar y localizar explosiones nucleares, también correrá a cargo de instrumentos transportados en satélites GPS.

Aunque los satélites militares de Estados Unidos y de la Unión Soviética tienen encomendadas misiones muy similares, presentan entre sí diferencias que afectan a su eventual vulnerabilidad ante las armas antisatélites. Estados Unidos ha ido siempre por delante en el campo de la microelectrónica y en otras tecnologías avanzadas, de ahí que los satélites norteamericanos sean mucho más refinados, más seguros y de más larga vida que sus adversarios soviéticos. La longevidad reviste particular importancia para los satélites de grandes altitudes, por los elevados costos a que obliga situarlos en las mismas. Por esa razón, Estados Unidos se ha esforzado en abarcar, con un menor número de redes de satélites, una amplia diversidad de funciones y en sacarles mayor rendimiento con sustituciones menos frecuentes. Aunque algunos satélites militares norteamericanos están aún situados en bajas órbitas, lo que los hace potencialmente vulnerables al actual sistema antisatélite soviético, la mayoría de los satélites de Estados Unidos se mueven en órbitas fuera de su alcance.



1. ARMA ANTISATELITE NORTEAMERICANA. Se ha diseñado para lanzarla al espacio desde un caza F-15 en vuelo a gran altura. El cohete que transporta el vehículo de aproximación miniaturizado aparece en esta fotografía en uno de los primeros ensayos, realizado para demostrar la compatibilidad del misil con el avión nodriza que lo alberga. La enseña que se aprecia en la cola del aparato indica la misión contra satélites del sistema de armas.

Por su parte, los satélites militares de la Unión Soviética ocupan fundamentalmente órbitas bajas; tienen, asimismo, una vida más breve. Probablemente esta circunstancia no sólo sea exponente de opciones operativas, sino también de deficiencias tecnológicas. Este último factor lo sugieren informes de que se produjeron fallos que impidieron a los soviéticos controlar las fases críticas del reciente conflicto de las Malvinas, a pesar de la gran atención que han venido prestando a los satélites de vigilancia oceánica. En cuanto a su vulnerabilidad ante las armas antisatélites, gran parte de la flota soviética de satélites se ve amenazada por la generación actual de armas antisatélites de baja altura, aunque el fuerte ritmo a que la Unión Soviética lanza satélites sustitutivos obvia en cierta medida ese punto débil.

Funcionarios y publicaciones del Departamento de Defensa de Estados Unidos han afirmado reiteradamente que la mayor frecuencia de los lanzamientos de satélites soviéticos demuestra que el programa espacial militar de la Unión Soviética es mucho más amplio que el norteamericano. Pero la disparidad de dichos intervalos entre una y otra potencia ha de ser atribuida a la más corta vida de los satélites soviéticos. De igual modo, muchas de las comparaciones que se hacen del peso bruto puesto en órbita ignoran la lanzadera espacial norteamericana. Con sólo dos vuelos de la lanzadera se sobrepasan los 300.000 kilogramos que viene poniendo en órbita cada año la Unión Soviética. También hay que tener gran precaución al considerar las comparaciones basadas en los gastos que se dedican a ambos programas espaciales. Con un razonamiento así se llegaría a la conclusión de que la Unión Soviética produce más alimentos que Estados Unidos.

Los satélites militares desempeñan papeles muy diferentes en tiempos de paz, en caso de pequeños conflictos o en una confrontación nuclear. En tiempos de paz, los satélites de baja órbita de fotorreconocimiento, de vigilancia oceánica y de información electrónica proporcionan datos fundamentales para el control de armamentos y para la información militar rutinaria. Si sobreviniese una crisis, estos satélites serían de gran valor para mantenerla controlada, pero tan pronto como comenzasen las hostilidades no nucleares asumirían un doble papel, al reforzar la eficacia de los combatientes. Llegado el momento de que una crisis política o un pequeño conflicto armado corriesen el

PRUEBA	FECHA	INTERCEPTOR (NUMERO COSMOS)	ALTITUD DE INTERCEPCION (KILOMETROS)	ORBITAS PARA INTERCEPCION	RESULTADO DE LAS PRUEBAS
1	20 OCT. 1968	249	525	2	FRACASO
2	1 NOV. 1968	252	535	2	EXITO
3	23 OCT. 1970	374	530	2	FRACASO
4	30 OCT. 1970	375	535	2	EXITO
5	25 FEB. 1971	397	585	2	EXITO
6	4 ABR. 1971	404	1,005	2	EXITO
7	3 DIC. 1971	462	230	2	EXITO
8	16 FEB. 1976	804	575	1	FRACASO
9	13 ABR. 1976	814	590	1	EXITO
10	21 JUL. 1976	843	1,630	2	FRACASO
11	27 DIC. 1976	886	570	2	FRACASO
12	23 MAYO 1977	910	1,710	1	FRACASO
13	17 JUN. 1977	918	1,575	1	EXITO
14	26 OCT. 1977	961	150	2	EXITO
15	21 DIC. 1977	970	995	2	FRACASO
16	19 MAYO 1978	1,009	985	2	FRACASO
17	18 ABR. 1980	1,174	1,000	2	FRACASO
18	2 FEB. 1981	1,243	1,005	2	FRACASO
19	14 MAR. 1981	1,258	1,005	2	EXITO
20	18 JUN. 1982	1,379	1,005	2	FRACASO

2. RESULTADOS DE LAS PRUEBAS del sistema de armas antisatélites adoptado por la Unión Soviética, resumidos aquí. La proporción global de éxitos del sistema ha sido del 45 por ciento (nueve interceptaciones en 20 intentos). En las pruebas realizadas con el primitivo ingenio de aproximación dirigido por radar, el arma ha obtenido una proporción de éxitos del 50 por ciento en las interceptaciones en primera órbita y del 70 por ciento en las interceptaciones en segunda órbita, con una tasa neta de éxitos del 64 por ciento. La totalidad de las seis pruebas registradas con el sensor óptico/infrarrojo, más complejo, se ha saldado con fracasos (*líneas en color*). El ritmo de realización de las pruebas y las órbitas escogidas para el interceptor permiten deducir que el arma soviética antisatélite podría estar pensada, no para atacar satélites norteamericanos, sino chinos. Datos de Nicholas Johnson en *The Soviet Year in Space: 1982*.

peligro de provocar un enfrentamiento estratégico nuclear, los satélites de gran altitud cumplirían esa doble función: constituirían importantes factores de disuasión y serían bazas de especial valor para las fuerzas estratégicas si el enfrentamiento fuera inevitable.

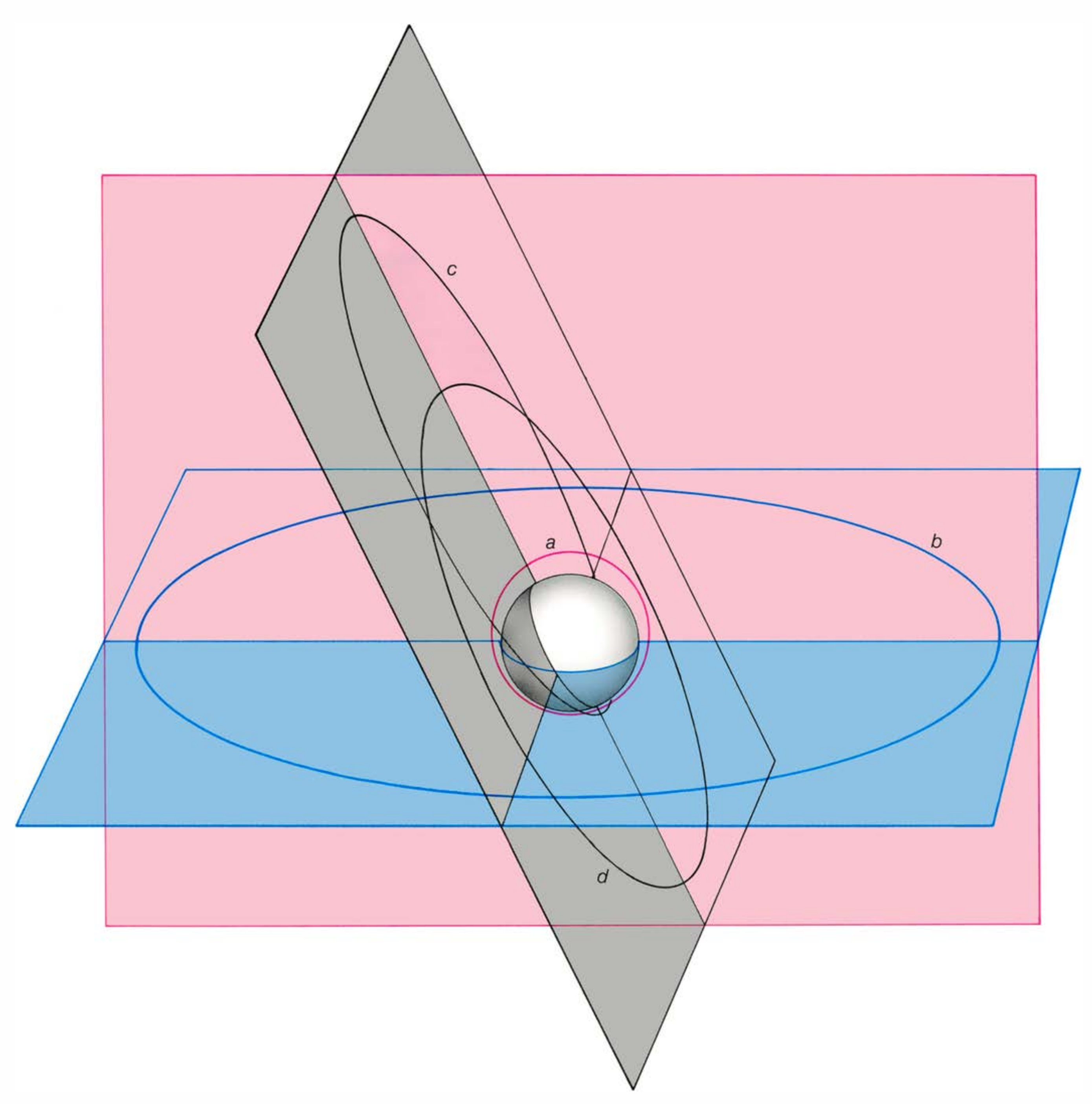
Todas las armas antisatélites desplegadas o con las que se están haciendo pruebas en tierra alcanzan una altitud máxima de algunos miles de kilómetros, si no menor. Por ello, solamente podrían atacar satélites situados en órbitas bajas o en órbitas muy elípticas. Dado que los satélites para la alerta precoz, la navegación, la evaluación del ataque y las comunicaciones, esenciales para las fuerzas estratégicas estadounidenses, se mueven en órbitas muy elevadas, no correrán ningún riesgo en un futuro próximo. La Unión Soviética se enfrenta, en cierto modo, a una amenaza potencial mayor: algunos de sus satélites esenciales para las comunicaciones, y todos los de alerta temprana, están situados en órbitas Molniya altamente elípticas.

¿Qué impacto puede tener la actual generación de armas antisatélites sobre la estabilidad de la crisis? En una guerra nuclear total, los sistemas instalados en el espacio y destinados a la dirección, al control, a las comunicaciones y la información (command, control, communications and intelligence; C^3I) se verían seriamente amenazados, existiesen o no armas antisatélites, puesto que las instalaciones para las comunicaciones situadas en el aire o en tierra y los centros de control se mostrarían sumamente vulnerables ante un ataque nuclear. Las partes del sistema C^3I que resultan esenciales para las fuerzas estratégicas de cada parte son hoy por hoy en gran medida inmunes a los ataques contra satélites; los satélites de baja órbita que sobrevivan a tales ataques serían de dudosa utilidad en medio de la gran destrucción que habría provocado en tierra el enfrentamiento nuclear. De todo ello se desprende que las capacidades contra satélites susceptibles de ser operativas en esta década tendrán escaso impacto sobre las posibilidades de las superpotencias para ganar una guerra nuclear total, sin que importe lo que entendamos exactamente por "total".

Sin embargo, la capacidad de respuesta contra satélites a baja altura podría incidir de una manera poderosa sobre las perspectivas de contención de conflictos armados de escala limitada, incluidas las hostilidades en las que pudiese verse envuelto un pequeño número de armas nucleares. Los satélites de baja órbita van alcanzando progresiva-

mente un grado de complejidad cada vez mayor, al igual que va aumentando su capacidad de proporcionar en cualquier circunstancia meteorológica información precisa, de una calidad y cobertura sin precedentes. Ni con buques, ni aviones o misiles podrá conseguirse el adecuado sustitutivo de los sistemas desplegados en el espacio. De esa capacidad de destruir satélites militares de baja órbita y del temor de que el adversario ataque en cualquier momento los satélites propios podría, en consecuencia, derivarse una irreprimible tentación de hacer desaparecer los satélites enemigos. Por tanto, esa capacidad de destruir con rapidez los satélites de baja órbita podría alimentar el fuego de una crisis política o de un conflicto a pequeña escala que, de no existir armas antisatélites, se hubieran resuelto diplomáticamente. Puesto que comúnmente se cree que lo más probable, en caso de guerra nuclear, es que ésta comience por un conflicto de pequeña escala, parece claro que la actual generación de armas antisatélites es algo más que un pequeño peligro en un mundo ya de por sí muy amenazado.

Un arma contra satélite ataca de muy diversos modos a su blanco. El actual arma soviética antisatélite es

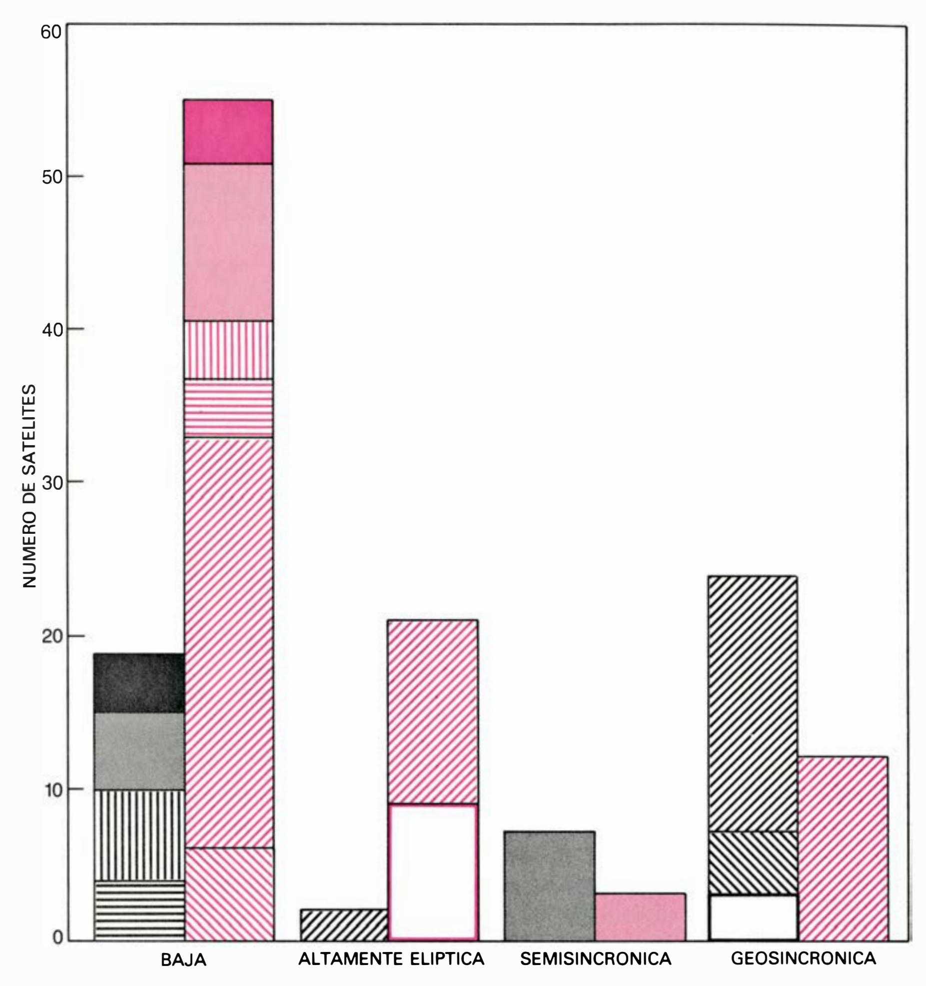


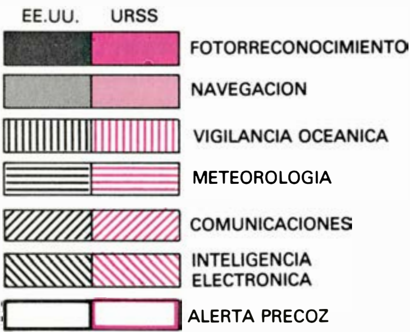
3. ORBITAS TIPICAS que describen los satélites militares de Estados Unidos y de la Unión Soviética. Se han representado a escala en un juego de planos que se cortan y permiten distinguir sus características. Los satélites cuya misión les exige moverse en órbitas relativamente bajas (entre 150 y 2000 kilómetros) tienden a ser más o menos circulares y polares, con un período de unos 100 minutos y una inclinación de entre 65 y 115 grados con respecto al ecuador (*a*). Los satélites de órbita geosincrónica giran alrededor de la tierra a unos 36.000 kilómetros de altitud y permanecen en un punto fijo sobre el ecuador (*b*). Los satélites de órbita altamente elíptica, con una inclinación típica de unos 63 grados, descienden en su punto más bajo hasta unos pocos centenares de kilómetros de altura sobre el hemisferio sur y alcanzan su cénit a unos 40.000 kilómetros, sobre el hemisferio norte (*c*). Los satélites en órbita semisincrónica suelen seguir una ruta prácticamente circular, a unos 20.000 kilómetros de altura y con una inclinación de entre 63 y 65 grados (*d*).

un interceptor coorbital: se le lanza hacia una órbita semejante a la de su objetivo y, cuando se acerca a éste, explosiona en una nube de metralla. El arma antisatélite norteamericana es un interceptor de ascensión directa: se le proyecta hacia el camino del satélite que constituye su objetivo por medio de un cohete, que un avión carga a su vez hasta una gran altitud. La destrucción se produce por impacto directo. No hay explosión.

Se han anticipado otras armas contra satélites; por su especial importancia, merece la pena detenerse en dos de ellas. La primera es una mina espacial: un pequeño satélite que, armado de una carga explosiva, acompañaría a su víctima en potencia durante semanas o meses, y que se dispararía por control remoto. La segunda es un sistema láser, que puede instalarse en tierra, en la atmósfera o en el espacio.

Ninguna de esas armas es nuclear. Tanto Estados Unidos como la Unión Soviética podrían también recurrir a misiles con carga nuclear como armas antisatélites, pero su utilización entrañaría graves perjuicios: se traspasaría el umbral nuclear en un momento crítico y se provocaría un impulso electromagnético que podría suponer un corte en el sistema de comunicaciones propio.





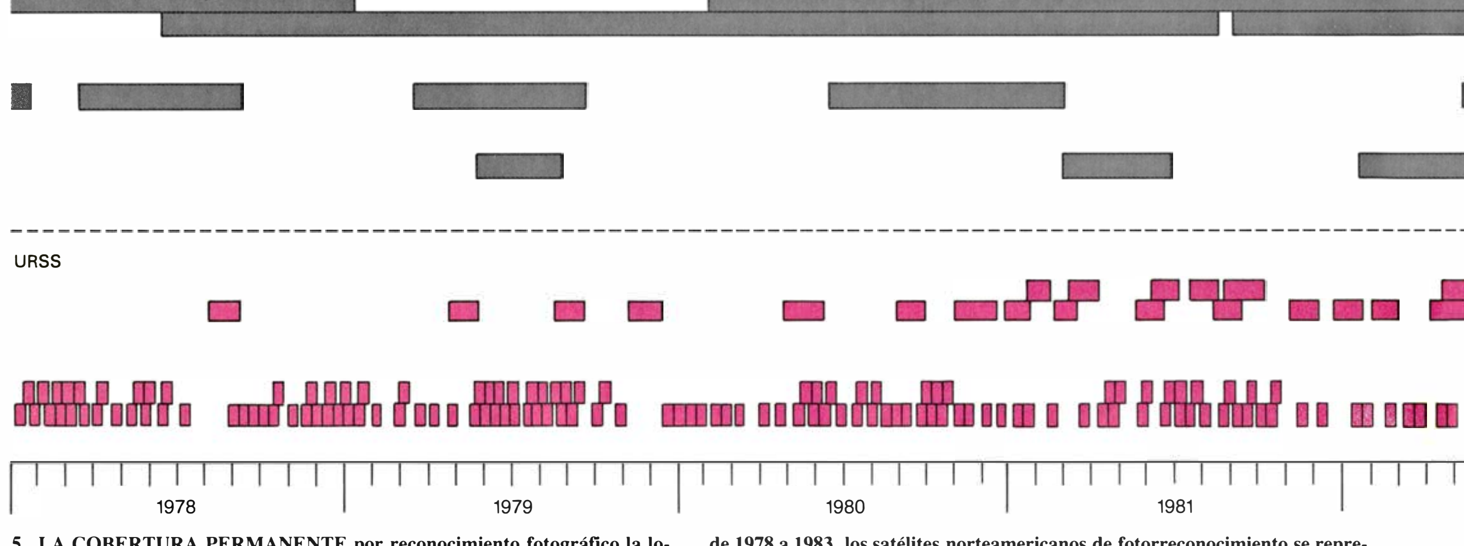
4. DISTRIBUCION TIPICA de los satélites militares desplegados por Estados Unidos y la Unión Soviética en un momento cualquiera, desglosada en las columnas de este cuadro, en razón de dos factores: altitud de la órbita y misión. Los únicos satélites potencialmente amenazados por la actual generación de armas antisatélites son los de órbita baja, representados en las dos primeras columnas.

La actual arma soviética antisatélite se introdujo en 1968. La pone en órbita un SS-9, un enorme misil balístico intercontinental de combustible líquido, cuyo despliegue empezó en los años sesenta. El interceptor pesa por sí solo más de 2000 kilogramos y mide unos seis metros. En todas las pruebas efectuadas, dicha arma antisatélite se ha lanzado desde Tyuratam, en la República de Kazakistán, hacia órbitas con una muy ligera diferencia de inclinaciones: entre 62 y 65 grados con respecto al ecuador. Desde Plesetsk, en el noroeste de la Unión Soviética, y hacia órbitas con la misma ligera diferencia de inclinaciones, se han lanzado satélites diana. La interceptación se produce una vez que el arma antisatélite ha dado una o dos vueltas en torno a la tierra, en un punto en el que su órbita se cruza con la de su objetivo. El arma antisatélite se acerca a su objetivo guiada por un dispositivo de aproximación que lleva a bordo: un radar activo o un sensor pasivo óptico o de infrarrojos.

La ausencia de variaciones de la inclinación orbital en el programa de pruebas soviético y el hecho de que existan tan pocos satélites norteamericanos con esas inclinaciones han llevado a especular si el arma soviética contra satélite no tendría como misión atacar satélites chinos más que satélites americanos. Existe, en efecto, un notable paralelismo entre los lanzamientos y las órbitas de los satélites chinos y el programa de pruebas soviético.

Cierto número de observadores cualificados (militares estadounidenses incluidos) han elaborado informes pormenorizados sobre el programa de pruebas soviético. Según esas fuentes, la Unión Soviética ha llevado a cabo una veintena de pruebas desde 1968 (la más reciente tuvo lugar el 18 de junio de 1982, como parte de un ejercicio estratégico que comprendía una diversidad de misiles balísticos y dos interceptores de misiles antibalísticos; la prueba antisatélite fracasó). La primera serie, que se realizó entre 1968 y 1971, se basaba en un radar activo para alcanzar el objetivo, después de dos órbitas. Desde 1976, la Unión Soviética ha explorado dos nuevas técnicas. Una de ellas permite un acercamiento más rápido con el dispositivo de aproximación dirigido por radar; la interceptación se intenta en la primera órbita. La otra busca la interceptación después de dos órbitas por medio de un dispositivo de aproximación óptico e infrarrojo. Este segundo sistema ha fallado en la totalidad de las seis pruebas, mientras que el acercamiento rápido por medio del dispositivo de aproximación de radar ha dado resultados positivos en dos de los cuatro intentos. La primera serie de pruebas, de 1968 a 1971, alcanzó resultados positivos en un 70 por ciento.

Aunque la versión soviética diseñada



5. LA COBERTURA PERMANENTE por reconocimiento fotográfico la logran de modo muy diferente Estados Unidos y la Unión Soviética. En general, los satélites militares norteamericanos son mucho más complejos, de más perfecto funcionamiento y de más larga vida que sus adversarios soviéticos, de lo que se deriva que Estados Unidos necesita muchos menos satélites para mantener el mismo grado de cobertura. En esta ilustración, que cubre el período de 1978 a 1983, los satélites norteamericanos de fotorreconocimiento se representan en gris y los soviéticos en color. El comienzo de cada barra señala la fecha en que se lanzó el satélite y su terminación señala la de su retorno a la tierra o la de su cese de operatividad. En la ilustración se incluyen todos los satélites de fotorreconocimiento de una y otra parte que han tenido una vida eficaz de tres o más semanas. La ilustración ha sido elaborada por los edito-

en 1968 parece ser bastante fiable, pueden oponérsele diversas contramedidas, como interferencias o desviaciones provocadas en el dispositivo de aproximación por radar o maniobras de evasión tan pronto se haya detectado el lanzamiento, pues la interceptación requiere unas tres horas. Las técnicas de aproximación y de maniobra efectuadas en los últimos ensayos han dado muy pobres resultados.

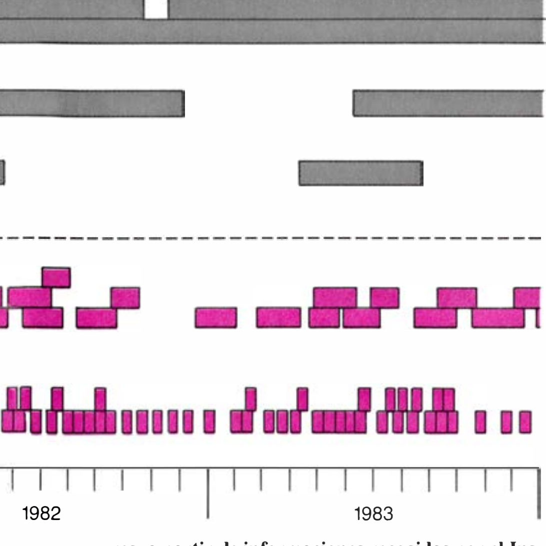
Aparte de estos defectos de funcionamiento, el sistema soviético contra satélites arrastra la rémora de imperfecciones en su concepción. En primer lugar, sólo puede atacarse un satélite si la ruta de su desplazamiento en tierra está próxima al lugar del lanzamiento del arma antisatélite. Esta condición se da en el caso único de que se trate de órbitas cuya inclinación sea más elevada que la latitud de dicho lugar de lanzamiento, lo cual a su vez sólo ocurre dos veces al día; y ello obliga a tener que esperar unas seis horas, por término medio, para intentar abatir el satélite en cuestión. En segundo lugar, la propia arma antisatélite necesita, por su gran tamaño y peso, un potente cohete que la impulse; sólo pueden lanzarse éstos desde un reducido número de instalaciones de las existentes en suelo soviético. En tercer lugar, no es fácil lanzar esos potentes cohetes impulsados con combustible líquido en una secuencia rápida desde un solo lugar de lanzamiento. Finalmente, y en cuarto lugar, la altitud máxima alcanzada en las pruebas soviéticas (unos 2400 kilómetros) está muy por debajo de las órbitas de los importantes satélites norteamericanos de navegación, alerta precoz y comunicaciones.

Considerados en bloque, esos factores permiten estimar que la fuerza contra satélites de la Unión Soviética tardaría por lo menos una semana en destruir los satélites norteamericanos a su alcance. Si se tratase de atacar satélites situados a 20.000 o más kilómetros, se precisaría un cohete impulsor aún más potente y, a juzgar por los resultados que ofrecen las pruebas realizadas, no parece probable que el arma antisatélite consiga una capacidad de maniobra y de dirección hacia su objetivo que le permitan interceptarlo a esas altitudes.

Puede concluirse, pues, que el actual sistema soviético contra satélites representa, para Estados Unidos, una amenaza que hay que matizar, nada flexible y bastante limitada. En términos similares se expresó ante el Senado de los Estados Unidos el Jefe del Estado Mayor del Aire, general Lew Allen, Jr., el 11 de julio de 1979: "Creo que la opinión general que tenemos es que hay que darles una muy dudosa capacidad operativa y para un reducido número de lanzamientos. En otras palabras, se trata de una amenaza que nos preocupa, pero no han desarrollado un programa de pruebas que pueda hacernos creer que se trata de una amenaza creíble". Nada importante se ha producido desde entonces en el programa de pruebas soviético que pueda hacer cambiar esta apreciación de 1979.

El arma antisatélite norteamericana de ascensión directa se lanza por medio de un cohete, cuyo reducido tamaño permite transportarla en un caza F-15. Este cohete de dos fases impulsa un instrumento, denominado vehículo de aproximación miniaturizado, hacia la trayectoria del satélite que constituye el objetivo, de acuerdo con las informaciones proporcionadas por la red terrestre de seguimiento de trayectorias de satélites. El vehículo de aproximación es un objeto cilíndrico que mide unos 30 centímetros de diámetro y pesa alrededor de 15 kilogramos. La aproximación se consigue mediante la acción combinada de ocho telescopios infrarrojos, un equipo de impulsores y un giroscopio láser. Los sensores infrarrojos perciben el objeto en el frío telón de fondo del espacio. La estabilidad se logra por rotación del cilindro; el giroscopio determina cuándo hay que poner en acción los impulsores para llevar el cilindro hasta la trayectoria de su objetivo. El choque se produce a gran velocidad y basta por sí solo para destruir el objetivo, aunque en la aproximación sólo se toleran errores minúsculos.

No se ha hecho público qué altura máxima puede alcanzar el arma antisatélite norteamericana. Pero, en una declaración ante el Congreso, se reveló que los objetivos prioritarios para el Estado Mayor Conjunto eran aquellos satélites cuyos objetivos fueran a su vez las fuerzas norteamericanas, presumiblemente satélites de fotorreconocimiento y de vigilancia oceánica, instalados en torno a los 500 kilómetros, o por debajo de esa cota.



res, a partir de informaciones recogidas por el Instituto Internacional de Estocolmo de Investigaciones para la Paz (SIPRI). Los autores del artículo no responden de la exactitud de los datos reflejados en ella y, por tanto, no son responsables de la misma. (Los dibujos son obra de George V. Kelvin.)

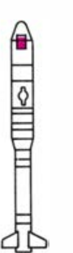
La primera prueba en vuelo del nuevo sistema antisatélite norteamericano tuvo lugar el 21 de enero de 1984. Según el Departamento de Defensa, un F-15 lanzó un cohete hacia un punto del espacio, pero sin llevar a bordo ningún vehículo de aproximación miniaturizado. Se prevé que en posteriores pruebas se incluya el propio interceptor de aproximación. El actual programa de pruebas se verá culminado, según se ha afirmado, con lanzamientos contra globos puestos en órbita a modo de objetivos. Está previsto que el sistema alcance capacidad operativa en 1987. Los planes originarios preveían el estacionamiento de dos escuadrones de F-15 especialmente equipados en las bases aéreas de McChord (Washington) y Langley (Virginia).

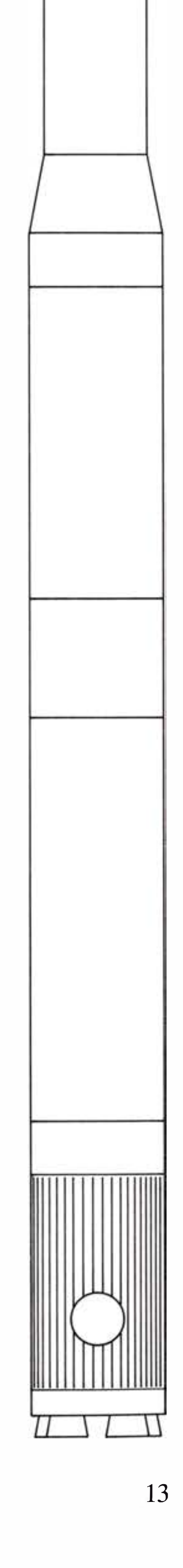
El tamaño notablemente reducido del vehículo de aproximación miniaturizado hace al sistema antisatélite norteamericano, en su completa explotación, mucho más versátil que el actual sistema soviético. En principio, puede adaptarse cualquier F-15 para transportar el arma antisatélite. Además, con aviones de transporte (o repostando en vuelo el F-15) podría conducirse el arma antisatélite hasta cualquier punto de ataque en el mundo. Si Estados Unidos realizase las necesarias inversiones en bases e instalaciones, podría alcanzar una fuerza capaz de destruir todos los satélites soviéticos de baja órbita en cuestión de horas. También podrían verse amenazados los satélites Molniya situados en órbitas altamente elípticas, pero es de suponer que la Unión Soviética situaría antes en órbitas más seguras sus satélites de alerta precoz y de comunicaciones.

Además de las armas antisatélites ya desplegadas o en vías de serlo, se esperan como mínimo otros dos tipos, para la próxima década. El ingenio más eficaz y de más fácil desarrollo parece ser la mina espacial. En principio, podría suponer una amenaza para todo tipo de satélites en órbita. Podría mantenerse estacionada durante largos períodos, siempre aparcada tan cerca que su detonación destruyera a la víctima. Podría incluso hacerse explotar por contacto. Pero el grado de refinamiento de una mina espacial dependerá del tipo de satélite al que acompañe. Un satélite de escasa capacidad de maniobra podría destruirse con una mina espacial bastante simple. Si el satélite que constituye el objetivo lleva sensores a bordo y puede maniobrar eficazmente, la mina tendrá que ser un ingenio refinado, tal vez más complejo que los actuales interceptores antisatélites, pero resultará siempre más barata que su blanco.

Son también de prever armas antisatélites que exploten rayos de "energía dirigida". En efecto, la Agencia de Proyectos de Investigación Avanzada para la Defensa ha lanzado un programa de pruebas cuyos dos principales componentes son un láser químico infrarrojo de dos megawatt y un refinado sistema de seguimiento y señalización, que podrían constituir, juntas, la base de un sistema antisatélite aerotransportado. Otras ideas con las que se especula en el campo de la energía dirigida, con la vista puesta en la defensa por misiles balísticos con base en el espacio, podrían también tener importantes posibilidades en un futuro más lejano.

Tras esta panorámica de la tecnología contrasatélite, podemos ya someter a valoración las amenazas que en el futuro representarán esas armas en un amplio abanico de hipótesis. Por mor de claridad, vamos a orientar nuestra





6. COHETES empleados para lanzar las armas antisatélites actuales de Estados Unidos y de la Unión Soviética; aparecen aquí en términos comparativos. El cohete norteamericano de dos fases, que se lanzó por vez primera desde un caza F-15 el pasado mes de enero, mide alrededor de cinco metros; la carga que transporta (el vehículo de aproximación no explosivo) mide unos 30 centímetros y pesa unos 15 kilogramos. El soviético, de tres fases, es un misil SS-9 modificado, de combustible líquido, cuyo despliegue comenzó en los años 60 para ser luego retirado de las fuerzas nucleares estratégicas. Mide unos 45 metros de longitud; la carga que transporta (la ojiva explosiva contrasatélites y su instrumental de maniobra) abarca unos seis metros, y pesa más de 2000 kilogramos.

visión hacia dos futuros diferentes: suponemos el primero sin ningún tipo de acuerdo negociado de limitación del desarrollo de armas antisatélites; el segundo, un futuro en el que estarían prohibidas todas las pruebas en vuelo de armas antisatélites.

Dadas las características del sistema antisatélite norteamericano, en un escenario de ausencia de limitaciones es muy verosímil que a finales de la década todos los satélites de baja órbita estén amenazados de destrucción rápida y simultánea. Para entonces, podría haber minas espaciales siguiendo a los satélites geosincrónicos y existir armas antisatélites dotadas de láser para atacar a los satélites de gran altitud.

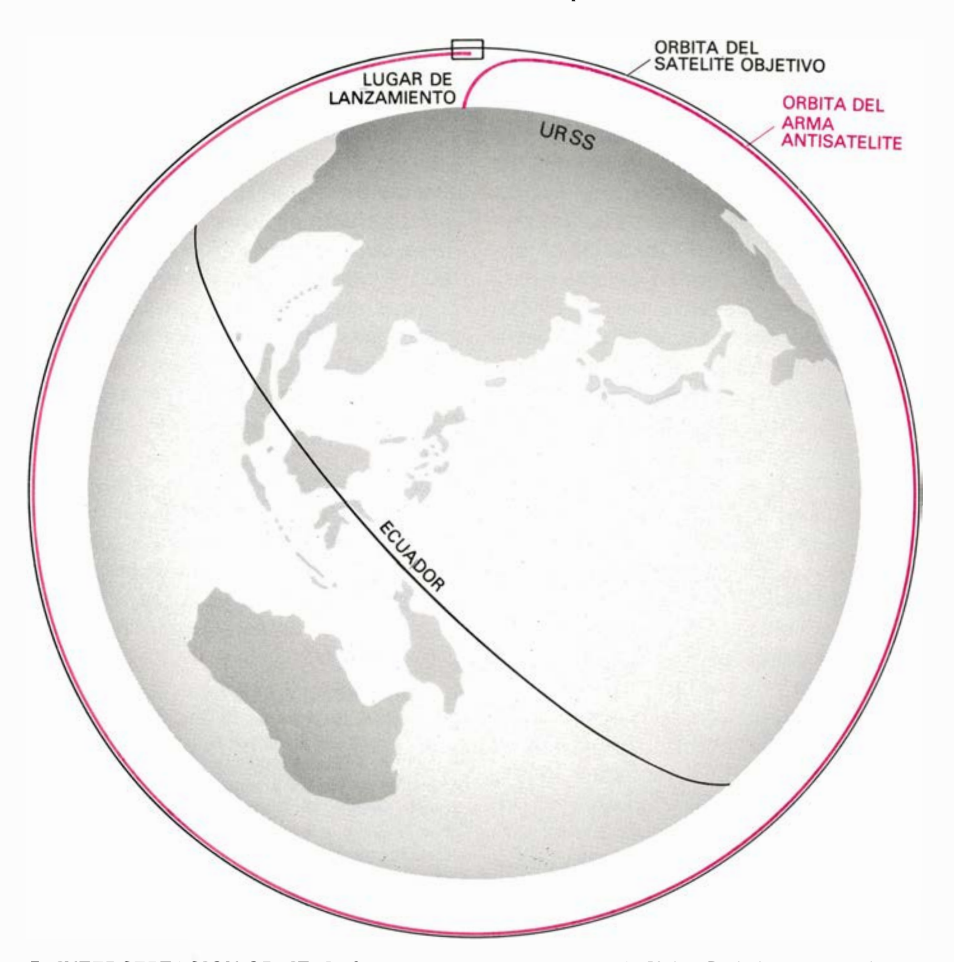
Paralelamente al aumento de las amenazas contra los satélites se dedicaría un esfuerzo cada vez mayor a la protección de los mismos, dotándolos de equipos de a bordo para evaluar las amenazas, aparatos para provocar interferencias en las comunicaciones, despliegues de trampas y cebos e impulsores aún más potentes, equipos todos ellos que les permitirían sobrevivir a algunos ataques, aunque la protección contra rayos láser intensos sería, en el mejor de los casos, difícil. Pero, sobre todo, habría que diversificar las funciones de los satélites, de modo que no podría seguir Estados Unidos confiando plenamente en un reducido número de muy complejos satélites multiusos. Por las mismas razones, habría que estar en posición de remplazar rápidamente los satélites de especial valor. Un plan tan complicado como ese sería muy costoso, si no inútil. Un satélite con una misión de vigilancia sutil o comunicaciones es intrínsecamente frágil y caro; nunca competirá con un ingenio más simple y barato, cuya única misión es destruirlo.

Si la fuerza antisatélite se convierte en una amenaza apremiante, habría que depositar mayor confianza en sistemas de apoyo: cohetes de sondeo e instalaciones C^3I. Avanzar en esa dirección supondría una desventaja para Estados Unidos, país que tiene sus fuerzas militares repartidas por todo el mundo, cuya seguridad descansa, de modo muy especial, en sistemas instalados en el espacio. (Alrededor del 70 por ciento de las comunicaciones militares norteamericanas a larga distancia se efectúan por satélite.) Por su parte, la Unión Soviética, potencia fundamentalmente continental, tiene una compleja red de comunicaciones instaladas en tierra y en el espacio. En opinión de algunos expertos, entre ellos Stephen M. Meyer, del Instituto de Tecnología de Massachusetts, los soviéticos consideran sus satélites de comunicaciones un apoyo más de su sistema terrestre.

Volviendo a las implicaciones políticas y militares de la carrera, ya hemos señalado que una amenaza real contra los satélites reduciría las posibilidades de que pudiesen mitigarse crisis y conflictos de todo nivel. Disponer de fuerza antisatélite contra satélites geosincrónicos tendría aún consecuencias más graves, pues aumentarían las posibilidades de un conflicto estratégico y podría darse al traste con algunos elementos esenciales del sistema que controlaría una guerra nuclear estratégica. Además de esos riesgos, la competencia en armamentos espaciales comportaría riesgos en sí misma. En una era de feroz lucha tecnológica, marcada por numerosas pruebas contra objetos en el espacio, hay que prever accidentes y sucesos poco claros, en los que se verán involucradas valiosas naves espaciales de uno o de ambos competidores. Existe, pues, un serio riesgo de que una carrera sin límites en armamentos espaciales pueda acarrear graves crisis, e incluso un conflicto armado.

En el segundo escenario, en el que estarían prohibidas las pruebas en vuelo de armas antisatélites, la Unión Soviética continuaría con su sistema de muy limitada capacidad contra satélites de baja órbita. Estados Unidos tendría un sistema potencialmente superior contra los mismos objetivos. Aunque aún se hallaría en fase de desarrollo, el sistema norteamericano podría desplegarse con operatividad en unos pocos años. En esa situación, Estados Unidos tendría las manos libres para mejorar la seguridad de sus satélites de baja órbita, diversificando sus funciones y aumentando su capacidad de maniobra y, si fuese preciso, podría instalar a bordo instrumentos destinados a provocar interferencias en radares. En esas circunstancias, la capacidad soviética contra satélites sería modesta, y sus posibilidades reales sufrirían un desgaste en ausencia de pruebas, mientras que la flota norteamericana de satélites se iría fortaleciendo progresivamente.

Cabe plantearse si contar con una capacidad contra satélites de baja órbita reviste una importancia tan apremiante para los militares estadouniden-



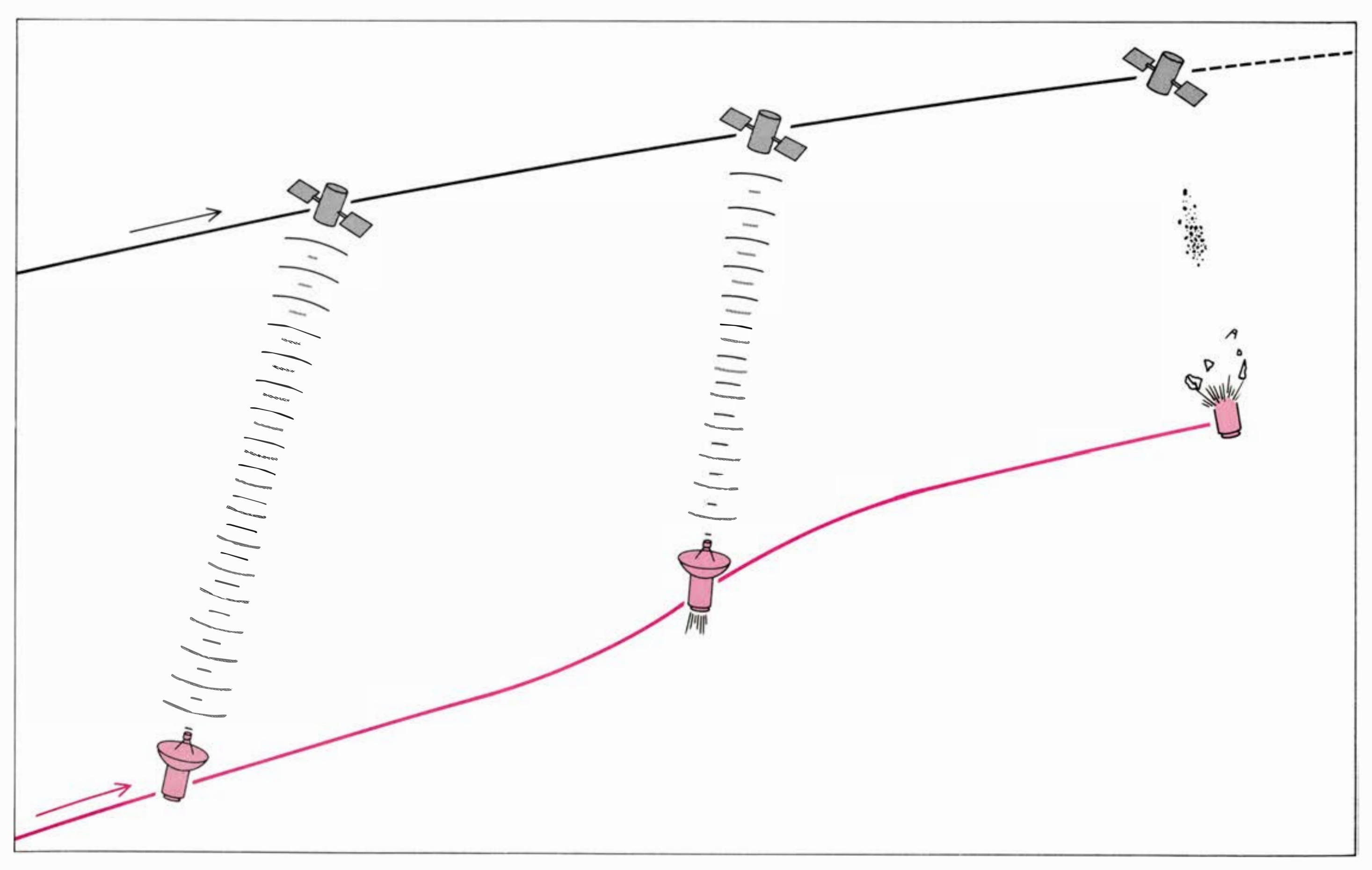
7. INTERCEPTACION ORBITAL, forma de ataque adoptada por la Unión Soviética para su sistema antisatélites. Se lanza el arma hacia una órbita prácticamente idéntica a la descrita por su objetivo, aunque un poco más alta o un poco más baja. En este caso, el arma antisatélite se halla en una órbita ligeramente inferior; es ella quien atrapa el objetivo, no a la inversa. En las fases finales del encuentro, el

ses, hasta el punto de hacerles ignorar los objetivos a largo plazo de estabilidad de la crisis y de la carrera de armamentos. Habrá que examinar, antes de contestar, qué objetivos se le señalan al arma norteamericana antisatélite. El primero lo constituye la propia arma soviética antisatélite. Pero no se trata de un objetivo verosímil: los soviéticos podrían desplegar, en las proximidades de su arma antisatélite y durante el breve tiempo en que permanece en el espacio, una serie de trampas en cadena que confundiría al interceptor norteamericano. Además, el interceptor soviético podría maniobrar para escapar del vehículo norteamericano que se aproximase.

Los satélites soviéticos de información electrónica y de reconocimiento por radar, que han venido preocupando a la marina estadounidense, constituyen objetivos más realistas. El almirante Noel Gayler, ex-comandante en jefe de las fuerzas de Estados Unidos en el Pacífico y antiguo director de la Oficina Nacional de Seguridad, acaba de publicar una evaluación comparativa de las ventajas e inconvenientes relativos que suponen las armas antisatélites para la marina. Gayler señala que es fácil provocar interferencias o engañar a un satélite radárico y que no puede rechazarse la información electrónica circunscribiéndola a las señales enviadas desde buques; de no adoptarse especiales precauciones, tales señales las detectan otros puestos de escucha que no precisan estar situados en el espacio. Sin embargo, pueden ya establecerse comunicaciones de alta densidad de datos entre buques y satélites por medio de rayos dirigidos con gran precisión, comunicaciones que son difíciles de interceptar. De acuerdo con esta opinión, la seguridad de la flota norteamericana de satélites reviste para la seguridad de la marina un interés mayor que la capacidad de destruir satélites soviéticos de vigilancia oceánica.

A medio camino entre los dos escenarios recién descritos (carrera ilimitada y prohibición de pruebas que limite ulteriores desarrollos) existe otra alternativa; a saber: permitir a ambas partes mantener y ensayar sus armas antisatélites de baja órbita y prohibir el desarrollo y las pruebas de cuantas armas abatieran satélites en órbitas geosincrónicas o en cualquier otro tipo de órbitas altas. En una situación semejante, los satélites de importancia para las operaciones militares convencionales correrían un riesgo, pero quedarían al abrigo los que se suponen imprescindibles para las fuerzas estratégicas.

Una pugna ilimitada en materia de armamentos contra satélites incidiría también sobre las armas estratégicas, habida cuenta de la estrecha relación que media entre éstas y la defensa contra misiles balísticos. Relación que se hace patente en el sistema norteamericano antisatélite, pues el vehículo miniaturizado de aproximación es, en esencia, el mismo ingenio diseñado para la interceptación en vuelo de ICBM por el programa para la realización de la aproximación (*Homing Overlay Program*) del ejército norteamericano. El director del laboratorio nacional de Los Alamos, Donald M. Kerr, Jr., acaba de firmar un reciente estudio donde examina con detalle la relación existente entre armas antisatélites y defensa de misiles balísticos. Implicación que abona George A. Keyworth II, consejero científico del Presidente, quien sugiere el desarrollo de un arma antisa-



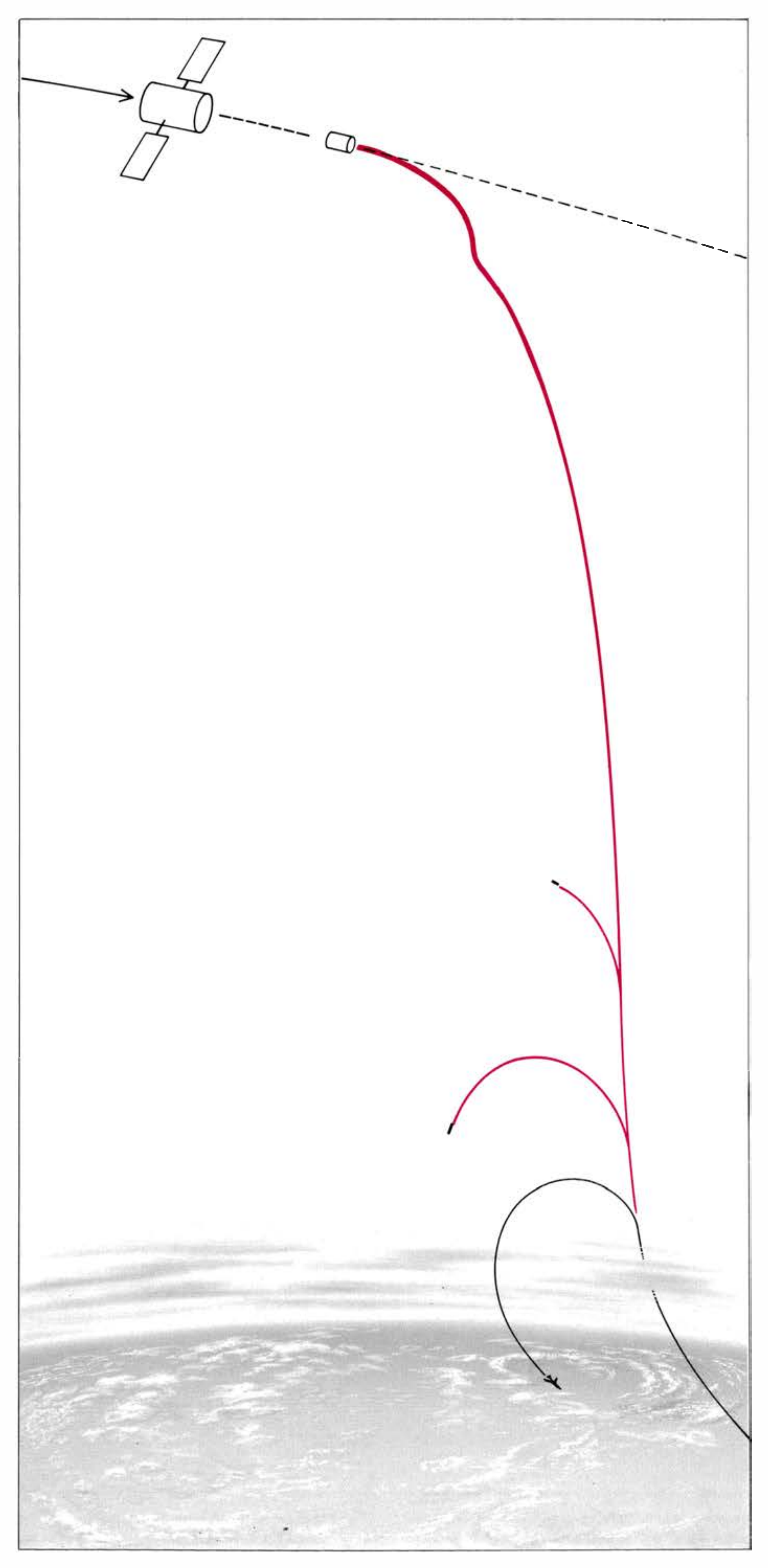
arma contra satélites y el objetivo se hallan en lo que puede denominarse órbitas de "caza y pastoreo". El arma antisatélites se acerca a su objetivo lateralmente, siguiéndole en su trayectoria, valiéndose para ello de un instrumento de aproximación óptico/infrarrojo dirigido por radar. Se produce entonces la explosión de la ojiva, que lanza una andanada de metralla contra el objetivo. En las 20 pruebas realizadas desde 1968 [*véase la figura 2*], el arma ha sido lanzada desde una base situada en la República de Kazakistán. La interceptación ha tenido lugar en la primera o en la segunda órbita.

télite láser como un primer paso hacia el logro de un sistema por misiles balísticos instalado en el espacio.

La relación tecnológica entre armas antisatélites y defensa de misiles balísticos comporta consecuencias de cierto relieve: ofrece una puerta por donde escaparse de los compromisos del Tratado sobre Misiles Antibalísticos (ABM) de 1972. Como señaló el presidente Reagan en su declaración de abril de 1983 sobre el alcance del control de armamentos, "la prohibición que establece el Tratado ABM de desarrollar, realizar pruebas y desplegar sistemas ABM instalados en el espacio, o componentes de tales sistemas, se refiere a tecnologías de energía dirigida o a cualquier otra tecnología usada a tales fines. Por ello, cuando tales programas de energía dirigida entran en la fase de pruebas en tierra, están sujetos a las limitaciones derivadas de las obligaciones del Tratado ABM". Pero el tratado ABM no prohíbe el desarrollo ni la realización de pruebas de tecnologías contra satélites. Sólo prohíbe la interferencia real con "medios tecnológicos nacionales" de comprobación, es decir, satélites usados con tales propósitos.

Mientras el Tratado ABM no se complete con otro que limite el desarrollo de armas contra satélite, cabrá siempre la posibilidad de pretextar que el arma que se está desarrollando pretende serlo contra satélite, aunque su fin último sea servir para la defensa de misiles balísticos. A diferencia de los satélites, los misiles son fuertes, pequeños y pueden quedar rápidamente rodeados de trampas; además, se mueven en trayectorias que hay que precisar al llegar el momento de la interceptación. En consecuencia, los sistemas defensivos de misiles balísticos se convertirían en eficaces armas antisatélites mucho antes de alcanzar su capacidad contra misiles balísticos. Será difícil refutar tal pretensión durante muchos años.

Ello es especialmente cierto en las armas de energía dirigida, que destrui-



8. INTERCEPTACION por ascensión directa. Es la planeada por Estados Unidos para su sistema antisatélites. El F-15 lanza el cohete de dos fases en dirección al objetivo, según la información de seguimiento recibida de una estación de control en tierra. El vehículo de aproximación miniaturizado alcanza la altitud orbital por medio de impulsores, que se separan, para caer uno tras otro, a medida que se acaba su combustible. El arma sigue entonces una trayectoria balística hasta que se aproxima a su objetivo, punto en el que el sistema de aproximación por infrarrojos toma las riendas y conduce el vehículo a la colisión con el objetivo. La destrucción se produce por impacto a alta velocidad. El actual programa de pruebas concluirá con lanzamientos contra globos-objetivo situados en órbita.

rían objetivos a gran distancia, y que por tanto representarían una amenaza potencial para los satélites geosincrónicos. Tales armas, en una defensa contra misiles balísticos, pueden hacer blanco en un minuto o antes, mientras que en el caso de satélites tardaríase horas. Por todas esas razones, un arma pensada exclusivamente con capacidad antisatélite inevitablemente habrá de contemplarla el adversario como un naciente sistema defensivo de misiles balísticos. Puesto que la forma más barata y eficaz de responder a una defensa contra misiles balísticos es la creación de fuerzas ofensivas estratégicas, ésta sería la consecuencia probable de un sistema innovador contrasatélite.

Un tratado sobre armas antisatélites podría abarcar uno o más de los objetivos siguientes: prohibir la realización de pruebas, así como poseer o usar tales armas; limitar las interferencias no destructoras de las funciones del satélite; prohibir cualquier clase de contacto con satélites susceptible de cambiar su trayectoria o su comportamiento o inutilizar su funcionamiento; imponer restricciones a todo tipo de acercamiento rápido y en proximidad entre dos objetos espaciales, sin importar de quién fueran, a fin de impedir la práctica de maniobras que pudieran considerarse contra satélites; institucionalizar un "código de circulación" que redujese en todo lo posible los encuentros fortuitos.

El análisis de todas estas medidas nos llevaría demasiado lejos. Vamos a concentrarnos, por el contrario, en aquellas disposiciones de los tratados que, desde nuestro punto de vista, serían fundamentales si de lo que se trata es de imponer limitaciones a la amenaza real contra los satélites, a través de una adecuada comprobación de su cumplimiento. A grandes rasgos, puede decirse que estos objetivos eran los que afirmaba perseguir Estados Unidos en las negociaciones suspendidas en 1979.

La forma más eficaz de impedir que aumenten las amenazas contra los satélites sería por medio de una prohibición de toda prueba de armas contra objetos espaciales, con independencia de dónde tuvieran sus bases las armas. Su comprobación se aseguraría renunciando a una prohibición de posesión de armas antisatélites, no poniendo ninguna restricción a las contramedidas contra satélites (mientras esas actividades no causen daño) y prohibiendo las pruebas de armas procedentes del espacio contra objetivos en tierra o en la atmósfera, pues tras esas actividades se pueden ocultar pruebas antisatélites. Naturalmente, el tratado prohibiría también el uso de armas contra satélites de otros países.

Dos de los autores (Garwin y Gottfried), acompañados del almirante Gayler, presentamos ante una sesión del Comité de Relaciones Exteriores del Senado de los Estados Unidos convocada al efecto un modelo de tratado inspirado en esas directrices que había elaborado un grupo seleccionado bajo los auspicios de la Unión de Científicos Involucrados.

El proyecto de tratado que presentó la Unión Soviética ante las Naciones Unidas en 1983 es semejante en los aspectos claves a nuestro tratado marco. El proyecto soviético de tratado pide también el rápido desmantelamiento de todas las armas antisatélites, pero no contempla la mayoría de los elementos del proyecto de tratado soviético de 1981 que criticamos en la sesión del Senado. En su última aparición en público, el presidente Andropov anunció una moratoria para las pruebas antisatélites por parte de la Unión Soviética, que se mantendría mientras Estados Unidos se abstuviese de realizarlas.

¿Podría comprobarse adecuadamente el cumplimiento de semejante prohibición de pruebas? Estados Unidos tiene por el planeta una red de radares y telescopios de seguimiento espacial que permite a sus fuerzas aéreas seguir todos los lanzamientos y objetos espaciales soviéticos. Dado que Estados Unidos tiene ya una experiencia de 15 años en el control de las pruebas soviéticas contra satélites no ocultas, las pruebas clandestinas del sistema actual sólo podrían escapar a su detección si lo que atacasen fuese un punto en el espacio, en lugar de un objetivo, y si evitasen realizar maniobras inusuales, característica de las armas antisatélites. A la vista de los resultados alcanzados por la Unión Soviética en este campo, es de pensar que no sería mucho lo que éste país ganaría con un programa de pruebas sometido a restricciones tan severas.

El control de las pruebas camufladas de nuevos tipos de arma antisatélites exige un examen caso por caso, pero todas esas pruebas se enfrentarán a un dispositivo de vigilancia cada vez mayor. La red norteamericana de radar se ha ampliado últimamente para darle mayor precisión en sus seguimientos y más capacidad de cobertura. El sistema de vigilancia electro-óptica espacial en profundidad con base en tierra (Ground-based Electro-Optical Deep Space Surveillance, GEODSS), que culminará en 1987, proporcionará datos ópticos y orbitales pormenorizados de objetos espaciales situados incluso en órbitas sincrónicas. Además, existiendo una prohibición de pruebas, puede tomarse la decisión de desarrollar vehículos espaciales diseñados expresamente para asegurar el cumplimiento de dicha prohibición. Tales vehículos podrían fotografiar y seguir la pista de objetos sospechosos, medir su emisión infrarroja para ver si sobre ellos inciden rayos láser y controlar aquellas actividades, en el espacio y en tierra, que, dentro de su legalidad, puedan enmascarar propósitos contra satélites.

Aunque una prohibición de pruebas impondría importantes obstáculos a los avances contra los satélites, dejaría las manos libres a ambas partes para proteger a sus propios satélites contra la amenaza residual que pudiera representar cualquier ingenio antisatélite aún existente. ¿Por qué no suprimir esta amenaza residual con la prohibición de la posesión de armas antisatélites, especialmente teniendo en cuenta la oferta soviética en ese sentido? Naturalmente, la prohibición sería de gran ayuda, pero hay que reconocer la dificultad de comprobar su cumplimiento. Incluso el interceptor soviético de mayor tamaño podría almacenarse en secreto, por no hablar del arma norteamericana, de tamaño menor. Además, si los soviéticos desmantelasen todas sus instalaciones para el lanzamiento de su arma antisatélite, pondrían en dificultades su propio programa espacial.

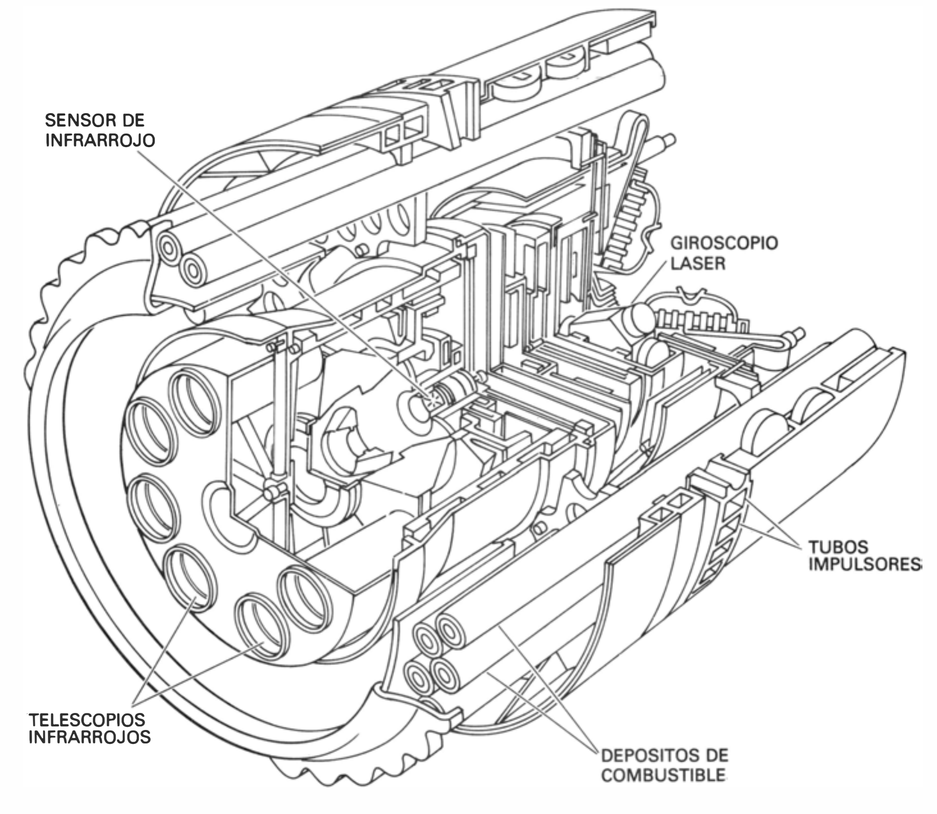
De ahí que la prohibición de poseer armas antisatélites y el desmantelamiento de las instalaciones utilizadas para las pruebas soviéticas contra satélites hayan de contemplarse como una valiosa medida para crear un clima de confianza, pero no como cláusulas comprobables de un primer tratado. En todo caso, si el funcionamiento del sistema soviético contra satélites se viese limitado a su nivel actual por una prohibición de pruebas, las afirmaciones de 1979 del Jefe del Estado Mayor de las Fuerzas Aéreas dejan bien claro que la Unión Soviética no podría obtener ninguna ventaja militar de importancia ocultando cierto número de sus interceptores.

La actitud de la Administración norteamericana ante el control de armas antisatélites quedó clara con el informe al Congreso del pasado 31 de

marzo. La Ley de Consignaciones para la Defensa para el ejercicio fiscal de 1984 contiene una cláusula donde se dispone de un fondo de 19 millones de dólares del que no se podrá hacer uso hasta que el presidente de la nación presente un informe que describa "los pasos concretos que se propone dar... en pro de un acuerdo probable con la Unión Soviética para prohibir o limitar estrictamente los sistemas ASAT, presentes y futuros".

La conclusión esencial del informe de la presidencia es que "no se ha podido llegar hasta ahora a ningún arreglo ni acuerdo, más allá de los que actualmente rigen las actividades militares en el espacio exterior, que pueda considerarse dentro de los intereses globales de Estados Unidos y sus aliados"; por tanto, no existe ningún propósito de emprender negociaciones en el momento actual. Varias son las razones que se citan en apoyo de esta conclusión: la posibilidad de que la Unión Soviética se guarde una remesa de interceptores antisatélites; la dificultad de controlar pruebas antisatélites; la consecuente amenaza de un repentino despliegue soviético de armas que se hayan desarrollado clandestinamente; la posibilidad de que armas no pensadas en principio con finalidades contra satélites (como los ICBM) se empleen para atacar a los satélites y a sus estaciones en tierra; el problema de cómo definir un arma espacial en el contexto de contramedidas electrónicas, naves espaciales y ciertos sistemas comerciales no especificados; y, por fin, la amenaza que representan para las fuerzas estadounidenses (sobre todo para la Marina) los satélites soviéticos de reconocimiento e identificación de objetivos.

Nos hemos detenido en los problemas planteados por la Administración estadounidense: el significado militar de los interceptores existentes que podrían permanecer en el arsenal soviético, la comprobación del cumplimiento de la prohibición de pruebas, la vulnerabilidad de la marina a los satélites soviéticos y la credibilidad de disuadir de cualquier ataque contra nuestros satélites con la amenaza de una acción de represalia sobre los de los soviéticos. No hemos entrado a discutir la vulnerabilidad de los satélites y su sistema de apoyo para el ataque con armas estratégicas, porque éste es un peligro que sólo puede mantenerse alejado permaneciendo en paz, así como tampoco nos hemos preocupado de amenazas indefinidas o marginales. Sin embargo, algunos puntos de éstos exigen, al menos, un somero comentario.



9. VEHICULO DE APROXIMACION MINIATURIZADO situado en el interior del satélite norteamericano. Se trata de un instrumento cilíndrico, pequeño y robusto, que incorpora ocho telescopios de rayos infrarrojos y 56 pequeños cohetes impulsores emplazados alrededor de su periferia. El giroscopio determina cuándo han de conectarse o detenerse los impulsores para ajustar la trayectoria del vehículo. El primer contrato para la fabricación del arma se ha concedido a la LTV Aerospace and Defense Company.

En primer lugar, la Administración afirma que no es posible comprobar el cumplimiento de una prohibición de pruebas. Esta idea se funda en la comparación entre la tecnología armamentista, en evolución, y la capacidad estática de comprobación, ignorando el hecho de que el gobierno de Estados Unidos ha manifestado tener un conocimiento detallado del actual programa soviético contra satélites. Un régimen de sometimiento a un tratado podría (en realidad, debería) ir acompañado de una adecuada inversión en más poderosas técnicas de vigilancia. En su declaración al Congreso, el primero de marzo, en apoyo del programa "Star Wars", el Subsecretario de Defensa para la Investigación y la Ingeniería, Richard D. DeLauer, expuso los ambiciosos planes para una instalación espacial de seguimiento y para alcanzar objetivos, que contaría con radar, láser y otros dispositivos. Se prevé que este vasto complejo pueda observar y seguir al mismo tiempo decenas de miles de objetos espaciales, reconocerlos y decidir automáticamente cómo atacarlos. "Tiene que funcionar adecuadamente, incluso en caso de producirse perturbaciones causadas por armas nucleares o por agresión directa del enemigo." Naturalmente, con un sistema más barato y sencillo podrían recogerse y analizarse los datos procedentes de hechos aislados y esporádicos que deben controlarse en tiempos de paz para que se mantenga la confianza en un tratado de prohibición de pruebas contra satélites, tarea que pueden llevar a cabo sin agobios hombres y máquinas.

En segundo lugar, hay que comparar el coste del mantenimiento de un determinado nivel de capacidad de los satélites, en presencia y en ausencia de armas antisatélites. Si se desarrollan y despliegan esas armas hay que contar con un sistema mucho más versátil y robusto y de más rápida actuación que si se tratase solamente de controlar una prohibición de pruebas. Si las armas ensayadas están legalmente dispuestas para atacarlos, los satélites habrán de contar con medidas protectoras mucho mejores y con un sistema de apoyo más perfecto que si se tratase de una situación en la que sólo se temiese la posible existencia de una amenaza antisatélite estrechamente limitada. Al evaluar los costes de un programa contra satélites,

no hay que olvidar los de esas sofisticadas instalaciones de vigilancia espacial, medidas protectoras y sistemas de apoyo.

En tercer lugar, la tesis de la Administración no distingue entre los varios sistemas que potencialmente representan una amenaza inmediata y simultánea contra todo lo que constituye una flota de satélites y sistemas que podrían atacar sólo a uno, o a lo más a unos pocos satélites, dando la alerta para una acción de huida o de sustitución. El interceptor norteamericano y las futuras armas de energía dirigida pertenecen a esa primera categoría, mientras que el interceptor soviético y los satélites comerciales y para maniobra de naves espaciales pertenecen a la segunda. La primera categoría representa una importante amenaza militar y es también la más fácil de limitar con un tratado que prohíba las pruebas.

Al juzgar el valor de toda propuesta de control de armamentos hay que reconocer que no hay tratado que pueda comprobarse en todos sus detalles. Lo que cuenta a la postre es si la actividad susceptible de escapar a la detección constituye un riesgo mayor que los peligros que entrañaría un régimen sin tratados. Con esta observación por delante, veamos cuál es el equilibrio entre los futuros alternativos que hemos postulado. El primer régimen, marcado por una carrera ilimitada, acabaría exponiendo a todos los satélites a la amenaza de una súbita destrucción; de ello derivaría un riesgo creciente de que la crisis y los conflictos escapasen al control. En el segundo régimen, limitado por un tratado de prohibición de pruebas, los satélites no estarían del todo a salvo, pero los peligros a que habrían de hacer frente se escalonarían, se irían identificando y podrían contrarrestarse con medidas de fácil arbitrio y de precio razonable.

Debe apoyarse el régimen no limitado sólo si una parte o la otra esperan alcanzar una ventaja duradera, en materia de armas, que les permita la destrucción de satélites y de misiles balísticos. Mas no son esas las perspectivas. La experiencia acumulada desde Hiroshima prueba que no hay tecnología en la que una superpotencia pueda mantener a largo plazo ventaja militar que represente una grave amenaza para la otra, y que el resultado de esa competencia es la reducción del nivel de seguridad para ambas partes. No hay prueba de que una carrera de armamentos espaciales sea la excepción.

Metalurgia superficial precolombina

Los orfebres de las culturas andinas sabían cómo cubrir el cobre con oro o plata y cómo tratar aleaciones de cobre, oro y plata de suerte que la superficie del metal estuviera formada exclusivamente de oro

Heather Lechtman

Cuando los conquistadores españoles fundieron los objetos de oro y plata que arrebataron a los incas como botín de guerra, se sorprendieron al descubrir que el metal obtenido era bastante impuro. Aunque por fuera los objetos parecían estar hechos de oro o plata, en realidad eran aleaciones de ambos metales con cobre. Al menos un milenio antes del predominio inca, los orfebres andinos habían desarrollado estas aleaciones junto con procedimientos para tratarlas de manera que los objetos, una vez acabados, presentaran una superficie de oro o plata puros. Los especialistas sabían asimismo cómo cubrir objetos hechos exclusivamente de cobre con una fina capa de oro o plata.

En el Laboratorio para la Investigación de Materiales Arqueológicos del Instituto de Tecnología de Massachusetts, mis colegas y yo hemos estado ocupados durante bastante tiempo en conocer la evolución de la metalurgia en América durante la época precolombina. Uno de los principales objetivos de la metalurgia en el Nuevo Mundo era la producción de metales y aleaciones que pudieran ser tratados especialmente de modo que mostraran colores muy apreciados en la cultura de sus productores. Además, en los últimos años hemos centrado nuestros esfuerzos en la descripción del complicado proceso que inventaron los orfebres andinos para dar color de oro y plata a objetos hechos con otros metales.

Por medio del análisis metalúrgico de artefactos de diversos períodos de la prehistoria andina, hemos identificado dos técnicas principales de coloración: recubrimiento por desplazamiento electroquímico (distinto del recubrimiento por electrólisis) y dorado o plateado por depleción (o agotamiento). Hemos reproducido ambos métodos de coloración en el laboratorio utilizando únicamente los materiales y técnicas que estuvieron al alcance de los especialistas de la época. En cada caso, los materiales y las técnicas a emplear vinieron determinados por el examen minucioso de los mismos artefactos. Los artesanos que inventaron y perfeccionaron estos métodos pertenecían a las culturas del área central andina.

De entre las altas culturas que existieron en la Sudamérica andina y conocidas por nosotros a través de la investigación arqueológica, fueron las más antiguas las que más contribuyeron al desarrollo de la tecnología metalúrgica. Estas son la cultura chavín (ca. 800-400 a. C.), que tenía su centro en las tierras altas del norte del Perú; la cultura moche (ca. 100 a. C.-800 d. C.), que se extendía por varios valles fluviales del desierto de la costa norte del Perú, y la cultura chimú (ca. 1150-1476 d. C.), el gran estado costero que se extendía desde el corazón del país de los moche hasta por lo menos la frontera con Ecuador. Las fuentes españolas indican que los incas se llevaron a los mejores especialistas en el trabajo del oro y la plata de Chimú a su capital imperial establecida en Cuzco.

Los orfebres de la cultura moche fueron sin duda los artesanos más experimentados de la tradición metalúrgica andina. Su virtuosismo y la calidad de sus obras no fueron igualados por ninguna de las culturas posteriores, incluida la cultura inca. Los dirigentes del estado inca dependieron en gran medida de las tecnologías metalúrgicas descubiertas por los moche, al recurrir a objetos con apariencia de oro y plata para mantener una situación de preeminencia política que les permitió gobernar un imperio que se extendía, a lo largo de unos 5000 kilómetros, desde Colombia hasta Argentina.

Recientemente se ha encontrado un número importante de objetos de metal de estilo moche en el extremo norte de la costa peruana, cerca de la frontera con Ecuador. El lugar del descubrimiento se conoce como Loma Negra; la mayoría de los objetos descubiertos están hechos de delgadas láminas de cobre batido. Fue el trabajo de láminas de metal, y no la fundición, la característica fundamental de la metalurgia andina. Para conseguir formas y figuras tridimensionales se unían piezas hechas de láminas de cobre trabajadas por separado. Cuando las piezas de Loma Negra fueron desenterradas estaban cubiertas con la característica capa verde formada por productos de la corrosión, principalmente óxidos y carbonatos de cobre, que se había formado a lo largo del tiempo que llevaban enterradas. Sin embargo, la eliminación de los productos de la corrosión mostró que el cobre había sido originalmente recubierto con una fina capa de oro o plata, de suerte que pudiera parecer que los objetos estaban hechos enteramente de esos metales preciosos.

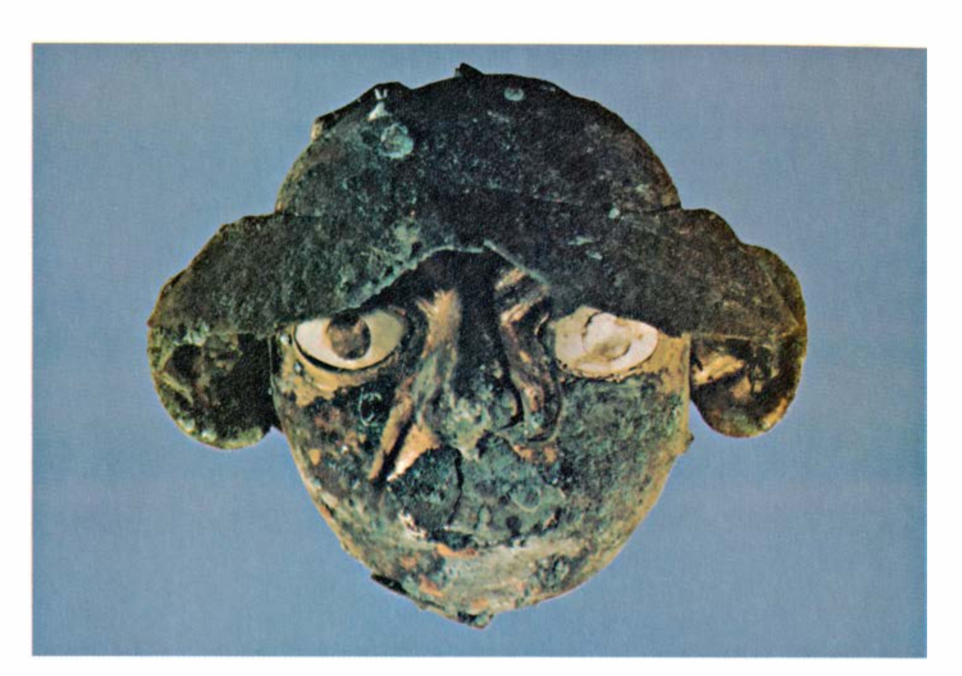
El examen metalográfico realizado en nuestro laboratorio de pequeñas muestras tomadas de algunas de las piezas de Loma Negra, reveló que muchos de los revestimientos superficiales eran tan finos –medían entre 0,5 y 2 micrometros de sección transversal– que resultaban casi invisibles al mirarlos con un aumento de 500 diámetros. Nuestros estudios nos llevaron a determinar las siguientes características de las superficies. Primero, el recubrimiento de

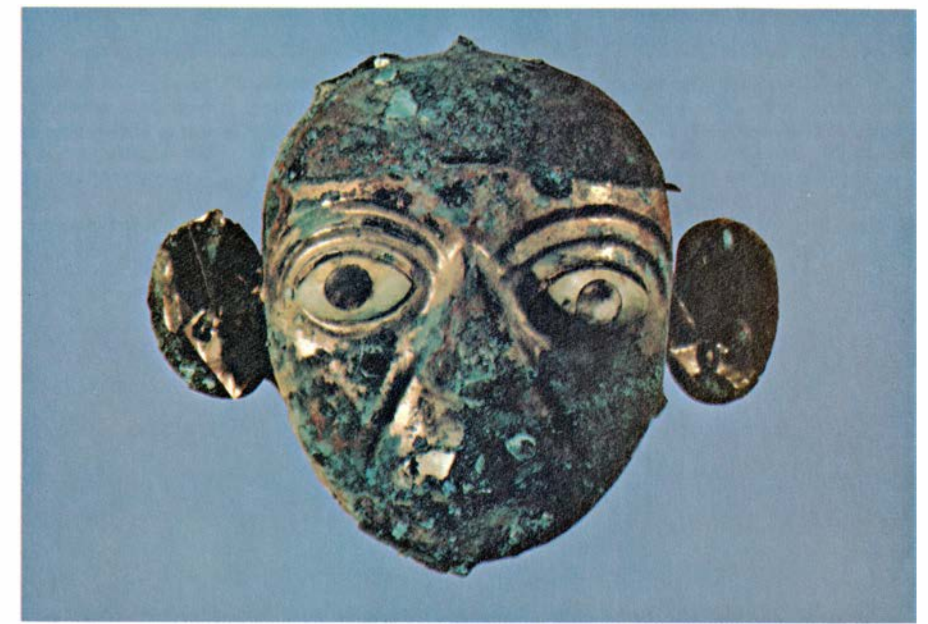
algunas de las piezas era de espesor extraordinariamente uniforme y cubría por igual toda la superficie, incluyendo filos tan finos como el papel. Segundo, una zona de difusión en estado sólido entre la capa superficial de metal precioso y la capa inferior de cobre indicaba que en alguna fase del proceso de recubrimiento se había aplicado calor a los objetos. Tercero, los artefactos de Loma Negra no mostraban signos de ninguna técnica de recubrimiento que fuera familiar para nosotros, tales como la amalgama, la aplicación de panes u hojas de oro, el baño de oro fundido, fusión de electrodos de oro o laminado Sheffield.

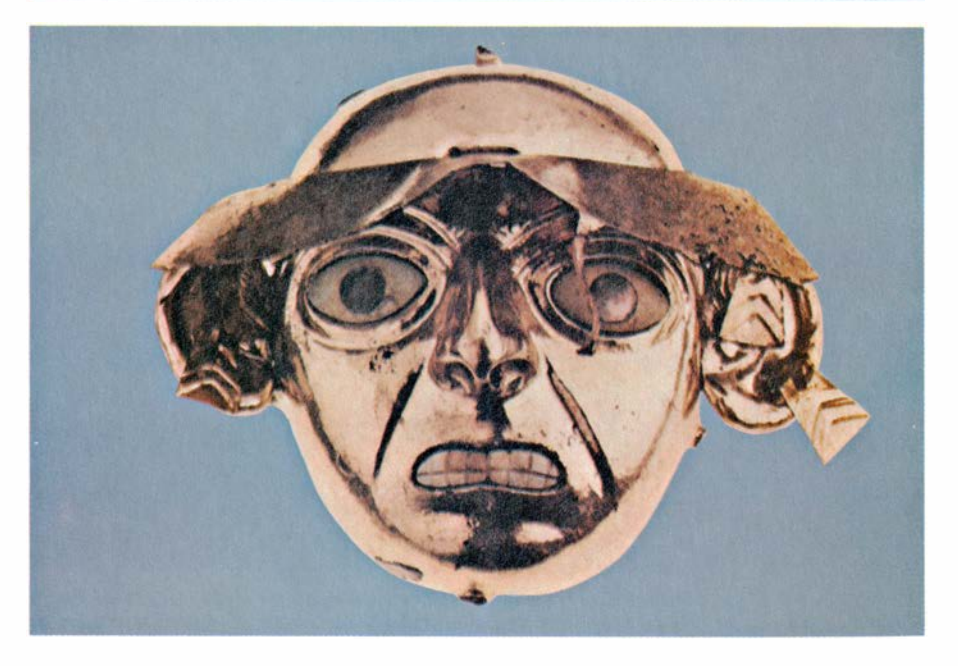
Los rasgos más sorprendentes de los recubrimientos de Loma Negra eran su extrema delgadez y uniformidad. Pero, al mismo tiempo, ambas características no pueden quizá ser el resultado de lo que más parecen: moderna galvanoplastia. En la galvanoplastia, los metales preciosos, presentes en forma de iones en un baño electrolítico, se depositan sobre los metales menos nobles sumergidos en el baño. Una fuente externa de corriente eléctrica, una batería por ejemplo, da lugar al proceso de recubrimiento mediante la reducción de los iones al estado metálico en el cátodo.

No obstante, la calidad de los revestimientos superficiales de los objetos de Loma Negra hacía pensar firmemente que habían sido depositados mediante algún método de recubrimiento químico. ¿Qué otras técnicas de deposición pudieron ser conocidas por los especialistas de la cultura moche? Una técnica de este tipo, de considerable antigüedad en el Viejo Mundo, es la de desplazamiento electroquímico. Los romanos la utilizaron para cubrir hierro con cobre y hacia el siglo noveno de la era cristiana se empleaba para preparar el hierro que debía ser dorado posteriormente. Los maestros armeros del siglo XVIII europeo utilizaron la misma técnica para decorar con oro las armaduras de hierro y acero.

El procedimiento empleado en Europa durante el siglo XVIII variaba según el metal que hubiera que dorar. Cuan-







1. MASCARA EN MINIATURA procedente de Loma Negra, un lugar de la cultura moche situado en el norte del Perú, que estaba cubierta de óxido de cobre después de permanecer enterrada durante cientos de años (*arriba*). Al quitar la visera se descubrió que la superficie que había bajo el óxido era de oro (*centro*). Una vez eliminado el óxido, la pieza brillaba intensamente (*abajo*). El oro se había aplicado por recubrimiento electroquímico.



2. MASCARA DE MOMIA de la cultura chimú, la mayor de este tipo. Mide 74 centímetros de ancho por 40 de altura. El lingote que fue batido para obtener la lámina de metal de que se hizo la máscara tenía un 40 por ciento de oro, 48 por ciento de plata y 12 por ciento de cobre. Este tipo de aleación se conoce con el nombre de tumbaga. La lámina fue batida, recocida (calentada para que recristalizara) y lavada (tratada químicamente para eliminar las escamas superficiales de óxido de cobre), en varias ocasiones. El resultado fue una superficie libre de cobre. La plata que había quedado en la capa superficial de plata y oro se eliminó con una pasta corrosiva. El oro que quedó en la capa superficial fue bruñido. El método es denominado dorado por depleción.

3. PAREJA DE PAJAROS procedente del área de la cultura quimbaya, en Colombia, realizada por fundición en vez de repujado; está hecha también de la aleación de oro, plata y cobre conocida como tumbaga. El cobre de la superficie original se disolvió utilizando quizás un ácido de jugo vegetal.

do se quería cubrir con oro un objeto de plata o de cobre, se preparaba una solución acuosa de oro disolviendo el metal en un baño de agua regia (una mezcla de ácido clorhídrico y ácido nítrico). Cuando se quería cubrir con oro el hierro o el acero, el oro se disolvía en una solución de sal común, salitre y alumbre. En cada caso, una vez que el oro estaba en solución, el baño era calentado a sequedad. El cloruro áurico ($AuCl_3$) era extraído con alcohol (o con éter en el siglo XIX) de la compleja mezcla de sales desecadas, y la solución alcohólica se aplicaba entonces al objeto que debía ser dorado. Pero ninguno de los ácidos destilados descubiertos en Europa entre los siglos XII y XVI (nítrico, sulfúrico y clorhídrico) y capaces de disolver el oro y la plata, eran conocidos por los metalúrgicos andinos.

Al mismo tiempo, los sectores desiertos de la costa peruana no carecen de minerales corrosivos que los metalúrgicos andinos pudieron conseguir sin esfuerzo alguno. Nosotros decidimos intentar la disolución de oro en soluciones acuosas de sales corrosivas que pudieran haber sido usadas por los moche. El método más simple y efectivo que encontramos fue preparar un baño corrosivo disolviendo en agua, a partes iguales, sulfato alumínico potásico, nitrato potásico y cloruro sódico (sal común). El oro calentado en esta disolución poco a poco durante un tiempo de dos a cinco días se disuelve con facilidad. (Incidentalmente, la mezcla contiene, entre otros iones, los que están presentes en el agua regia.) Pero la solución es altamente ácida y ataca inmediatamente el cobre sumergido en ella. Teníamos que neutralizarla antes para que el oro de la solución pudiera cementarse sobre el cobre. La sal más efectiva que encontramos para este propósito fue el bicarbonato sódico, y el pH óptimo (la medida estándar de acidez o alcalinidad) para que se produjera el recubrimiento resultó ser 9, esto es, algo alcalino.

Después de cinco minutos de ebullición suave, una lámina de cobre sumergida en la solución neutralizada resultó cubierta uniformemente de oro por toda su superficie, incluidos los bordes. La película de oro tenía un espesor aproximado de un micrometro. Pero el proceso de recubrimiento era insuficiente para unir permanentemente el oro al cobre. A menos que el cobre estuviera extremadamente limpio y libre de óxido, la capa de oro tendería a desprenderse.



4. REGIONES DEL NUEVO MUNDO en las que se practicaba el dorado y plateado por depleción (*color*). Los métodos de depleción se conocían en Perú en el segundo milenio a. C. Los andinos trabajaban fundamentalmente con láminas de metal. Los metalúrgicos de las poblaciones que se extendían desde el norte de Colombia hasta México fundían la tumbaga por el procedimiento de la cera perdida.

Nuestros estudios metalográficos de los objetos de Loma Negra revelaron una inequívoca zona de difusión de estado sólido entre la capa de oro y la superficie inferior de cobre, lo que mostraba que los especialistas moche habían calentado sus piezas de cobre chapado para conseguir una fuerte unión metalúrgica en esta interfase. Cuando nuestras muestras preparadas en el laboratorio fueron calentadas durante sólo unos cuantos segundos a una temperatura entre 650 y 800 grados Celsius, también nosotros conseguimos una unión excelente entre la capa de oro y el sustrato de cobre.

Por fin, nuestros experimentos tuvieron éxito al conseguir recubrimientos de oro por procedimiento electroquímico que se parecían bastante a los de los artefactos de Loma Negra, tanto en su aspecto exterior como en su microestructura. Estos resultados sugieren que los especialistas andinos consiguieron sus recubrimientos de una forma muy semejante, disolviendo oro y plata en una solución acuosa de minerales corrosivos, comunes en los desiertos de la costa peruana. Es posible que su familiaridad con estos minerales la hubieran adquirido en el transcurso del desarrollo de otras tecnologías andinas, tales como la tinción de tejidos. Cualquiera que sea el origen, parece que la utilización de corrosivos naturales fue un elemento dominante y muy antiguo



5. SUPERFICIE DE UNA LAMINA de metal recubierta con oro por procedimiento electroquímico en el laboratorio de la autora, en el Instituto de Tecnología de Massachusetts, y parcialmente bruñida después, ampliada 500 diámetros en una micrografía electrónica de barrido. Se utilizó cobre. El límite entre la superficie bruñida (*izquierda*) y la que no lo está (*derecha*) se encuentra en el centro de la micrografía.



6. LA SUPERFICIE NO BRUÑIDA DE LA LAMINA, aumentada 2000 diámetros en esta micrografía electrónica de barrido, aparece recubierta de poros con aspecto de cráteres. Los poros surgieron como consecuencia del proceso de recubrimiento electroquímico; cada poro ocupa el lugar de un foso anódico (cargado positivamente) en la superficie del cobre. El cobre se introduce en un electrolito en el que se ha disuelto oro. Este se deposita en las superficies próximas a los fosos anódicos; las superficies son catódicas (cargadas negativamente). Al cubrir la superficie, el oro forma un anillo alrededor de cada poro.

en la metalurgia superficial en los Andes.

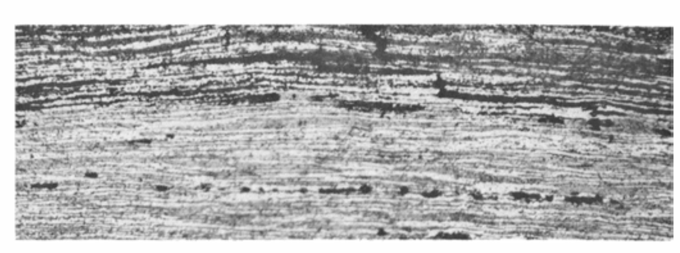
El recubrimiento por desplazamiento electrolítico se produce cuando un metal, cobre por ejemplo, situado en el extremo inferior o negativo de la serie electroquímica, se introduce en un baño electrolítico que contiene iones de un metal más noble, como el oro, situado en el extremo superior o positivo de dicha serie. Se produce una simple reacción de desplazamiento; el cobre desplaza al oro en la solución: $2AuCl_3 + 3Cu \longrightarrow 2Au + 3CuCl_2$.

Pero no nos revela nada del mecanismo más allá del mero proceso de recubrimiento. Ese mecanismo es exactamente igual que el de la más simple cuba utilizada por los modernos galvanizadores. Se requiere un área anódica (eléctricamente positiva) y otra catódica (eléctricamente negativa), ambas deben estar en contacto con el electrolito, y un circuito para que los electrones fluyan a través del metal desde el ánodo hasta el cátodo, equilibrado por el flujo contrario de iones a través del electrolito. A menos que la geometría del sistema sea tal que las áreas anódicas estén adecuadamente en contacto con el electrolito, solamente se formará un depósito infinitesimal antes de que el potencial de electrodo sea igual por todas partes y la acción recubridora se detenga.

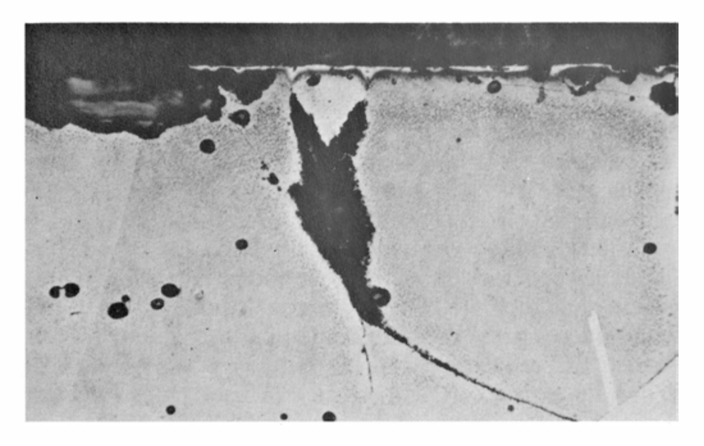
Cuando se recubre un metal por el método de desplazamiento, tanto el área del ánodo como la del cátodo están formadas por la superficie del mismo metal. En nuestros propios experimentos, como debe haber sucedido en el caso de los especialistas moche, agujeros o irregularidades de pequeño tamaño de la superficie de la lámina de cobre actuaron inicialmente como ánodos. Continuaron su actividad anódica hasta que quedaron completamente aislados del electrolito por el oro que se depositó sobre la superficie catódica adyacente. Las áreas catódicas y anódicas permanecen en un delicado equilibrio dependiente de la disponibilidad de iones, de la velocidad a la que los iones se difundan por el electrolito y del grado en que estén eléctricamente polarizadas las áreas. Sólo se obtiene un buen depósito cuando las áreas anódicas se han retraído hasta hacerse casi invisibles, pero no quedan completamente aisladas del electrolito.

La tradición andina de batir láminas de metal y unirlas para obtener objetos tridimensionales puede haber llevado al descubrimiento de otro método para

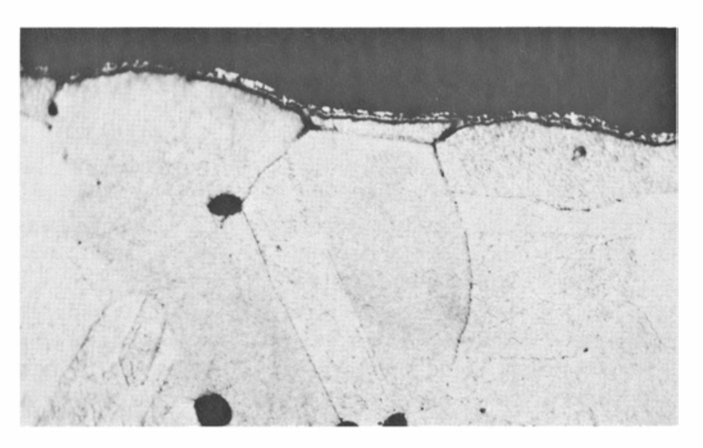




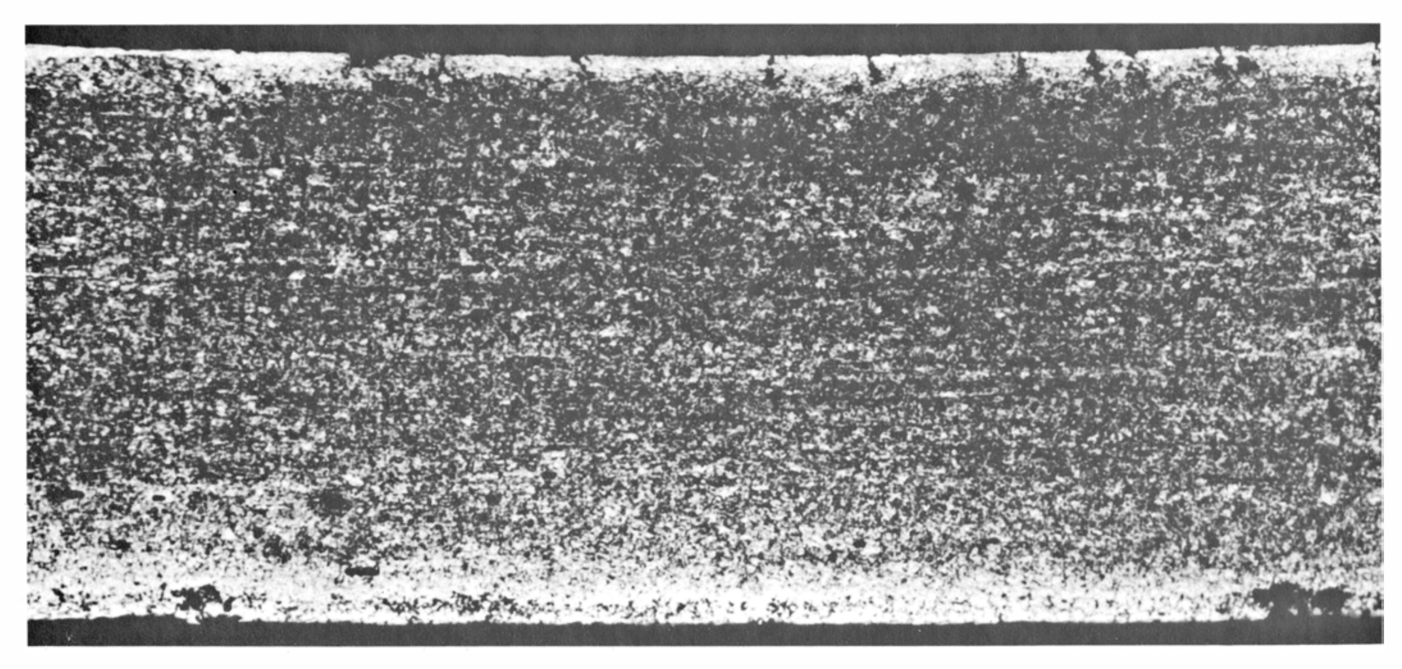
7. ABALORIO PERUANO procedente del yacimiento de Malpaso, en el valle de Lurín (*a la izquierda*). Está hecho de una aleación de cobre y plata. La capa superficial formada únicamente de plata se consiguió por medio de sucesivos batidos, recocidos (calentados para que recristalizaran) y lavados. La pieza mide sólo 20 milímetros y se trata del ejemplo más antiguo de aleación de cobre y plata que se conozca. La fotografía de la derecha muestra una micrografía de la sección transversal de la pieza con un aumento de 500 diámetros. Cuando el lingote del que se formó la lámina de metal se enfrió después de fundir la aleación, se separó microscópicamente en una fase rica en plata y un eutéctico doble formado por fases ricas en plata y en cobre. (Un eutéctico es la composición de una aleación que tiene el punto de fusión más bajo.) El subsecuente batido y recocido dio lugar a que las fases ricas en plata y en cobre se separaran casi completamente, como muestran las capas alternativas que se distinguen aquí. Las capas ricas en cobre están corroídas.



8. LAMINA DE COBRE DORADO procedente de una figura de Loma Negra vista en sección transversal en una micrografía con un aumento de 500 diámetros. La línea fina de color claro que aparece en la superficie es la capa uniforme de oro cementada por el recubrimiento electroquímico; en algunas zonas la capa llega a tener una altura mínima de 0,2 μ. La línea en *V* de la parte inferior derecha señala el límite de un grano de cristal de cobre. Los espesamientos en *V* mucho más pequeños situados en la misma capa de oro son consecuencia de la difusión del oro entre los límites de los granos de cobre cuando se calentó la pieza para unir ambos metales. Los circulitos oscuros son óxidos de cobre; las superficies negras más extensas, zonas de corrosión.



9. OTRA LAMINA DE COBRE que fue recubierta electroquímicamente en el laboratorio de la autora en el MIT sumergiéndola en una cuba electrolítica que contenía oro en solución. La cuba contenía una solución acuosa de minerales corrosivos de los que pudieron disponer los especialistas metalúrgicos de la cultura moche; esos minerales se encontraron en abundancia en los desiertos de la costa peruana. En esta micrografía óptica la superficie de la lámina se ha aumentado 1000 diámetros. Una vez recubierta la pieza, se recalentó durante 15 segundos con una llama. Como en el ejemplo precedente de Loma Negra, este calentamiento hizo que el oro se introdujera entre los bordes de los granos de cobre, uniendo firmemente de este modo ambos metales.



10. MUESTRA PROCEDENTE DE LA MASCARA CHIMU reproducida en la figura 2 vista en sección transversal en una micrografía con un aumento de 200 diámetros. Las dos superficies son aparentemente más ligeras que el interior de la muestra. Esto refleja la mayor riqueza en plata y oro de la superficie como consecuencia de que el proceso de dorado por depleción eliminó el cobre de la aleación al ser ésta sometida a diversos ciclos de recalentamiento y lavado. Las estrías visibles en la muestra representan fases de diferente composición que han ido alargándose durante los repetidos batidos de la lámina.

enriquecer la superficie con un metal precioso. El método, al que me referiré como dorado por depleción (*depletion gilding*) o plateado por depleción (*depletion silvering*), comenzó por la fundición de un lingote hecho no de cobre sino de una aleación de diversas proporciones de cobre y plata o cobre y oro, o los tres metales a la vez.

Una aleación de este tipo se comprime cuando se golpea con un martillo; como el lingote se va convirtiendo en una lámina, reduciéndose gradualmente su espesor, al final el metal acaba siendo demasiado duro y quebradizo para darle cualquier forma. Cuando se alcanza este punto, el orfebre debe recocer la pieza: poner el metal al rojo vivo para que recristalice y recupere su maleabilidad.

La más antigua de las aleaciones andinas que ha descubierto la arqueología es una aleación de cobre y plata. Su maleabilidad le permitía ser repujada y la dificultad que ofrecía al repujado garantizaba que, una vez terminado el trabajo, la lámina mantuviera la forma que se le dio. Cuando en el transcurso de la labor una lámina de aleación de cobre y plata es recalentada para que recupere su maleabilidad, el cobre de la superficie reacciona preferencialmente con el oxígeno del aire para formar una capa de escamas de óxido de cobre. La plata se oxida con menos facilidad; por lo tanto, después de las sucesivas secuencias de recalentado, desescamado, batido y nuevo recalentado que son necesarias para la manufactura de la lámina de metal, se forma gradualmente una capa superficial muy empobrecida en cobre y por consiguiente rica en plata.

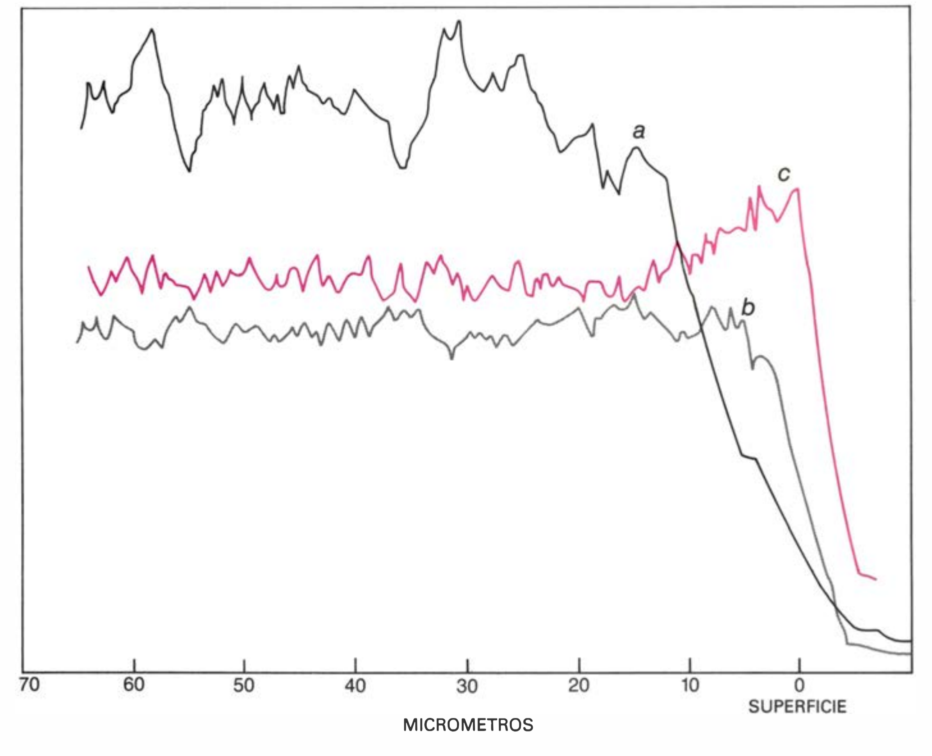
Lógicamente, después de varios ciclos como el descrito, la superficie de la lámina aparece como si fuera de plata, ya que es precisamente eso lo que ha quedado en ella. El óxido de cobre que se forma en el transcurso de cada recalentado tiene color marrón o negro. Para lavar la lámina, esto es, para eliminar las escamas por disolución, los orfebres andinos podrían haber aplicado orina rancia (la urea de la orina fresca se degrada en amoniaco) o los jugos ácidos de algunas plantas. En cada caso, la formación de una superficie plateada fue consecuencia ineludible de un calentamiento en el aire.

Se sabe que algunos objetos andinos hechos de aleación de plata y cobre son anteriores a la época de florecimiento de la cultura moche. Sin embargo, fueron los orfebres moche los que más sobresalieron en la utilización de estos materiales. Trabajaban con aleaciones con una concentración de plata que variaba dentro de una gama muy amplia: desde más del 50 por ciento de riqueza en plata hasta menos del 20 por ciento. Los objetos que manufacturaban eran duros y tenían al tacto y a la vista el aspecto de la plata. En tiempos posteriores a la cultura moche se utilizaron los mismos tipos de aleaciones por razones similares.

Pero la más importante de las aleaciones utilizadas y desarrolladas por los moche era una mezcla de cobre y oro conocida en el Caribe, donde los españoles aprendieron la palabra, como tumbaga. Cuando el oro y el cobre se funden juntos, se mezclan con facilidad. Después de enfriarse y solidificar permanecen mezclados; forman series de soluciones sólidas con toda la gama posible de proporciones. Las aleaciones de oro y cobre precolombinas variaban ampliamente en su composición: algunas eran tan bajas en oro que llegaban a tener sólo un 12 por ciento de este metal. El color de la aleación depende de su composición. Las tumbagas con una alta concentración de cobre son de color rojo o rosa; las que tienen mayor cantidad de oro son amarillentas.

También se encuentra plata en muchas de estas aleaciones. Esto puede ser debido a que el oro se añadiera a una aleación de plata y cobre o a que el oro ya contuviera cierta cantidad de plata, dado que las pepitas de oro de los placeres andinos la contienen normalmente. En cualquier caso, estuviera formada por oro y cobre o contuviera también plata, la aleación recibía el mismo nombre: tumbaga. Aunque la tumbaga se fabricó por primera vez en los Andes centrales, en épocas posteriores fue una aleación común en muchos otros lugares de América Central y del Sur en la época precolombina; fue conocida incluso en México.

El recalentamiento de una lámina de cobre y oro, de la misma forma que el recalentamiento de las láminas de cobre y plata, produce finalmente una superficie enriquecida en el metal precioso, ya que a través del proceso de oxidación se pierden cantidades cre-

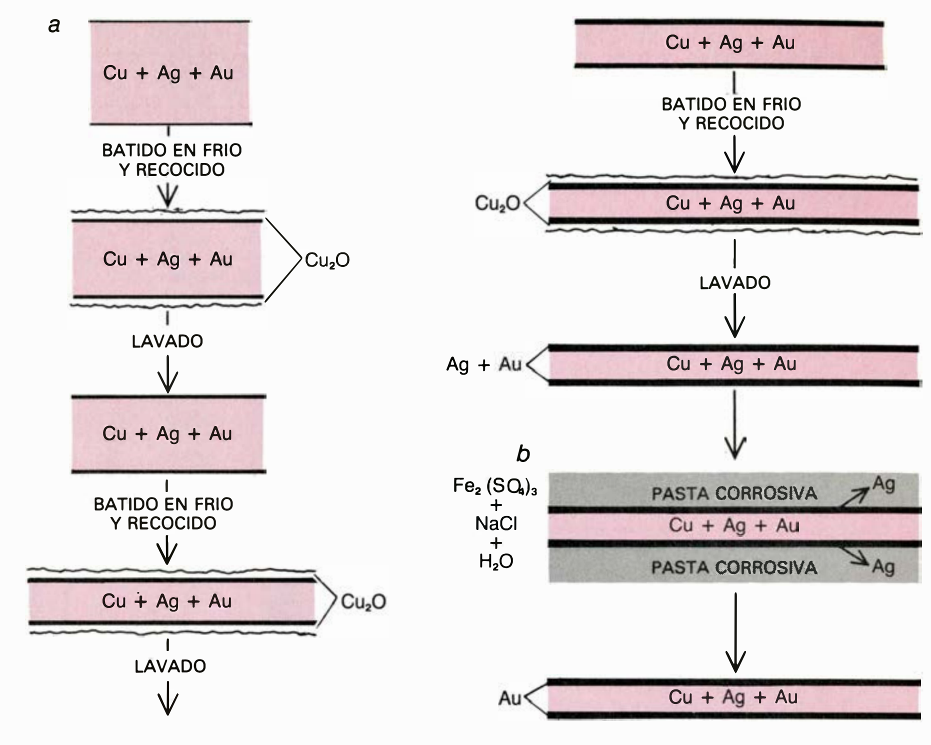


11. PERFILES DE CONCENTRACION del mismo ejemplo procedente de Chimú que muestra cómo la concentración de cada metal (oro, plata y cobre) cambia desde el centro de la sección transversal de la muestra hasta la superficie. La concentración en cobre (*línea negra*) permanece prácticamente uniforme hasta que el haz del análisis electrónico llega a una distancia de unos 12 micrometros de la superficie de la muestra (*a*). En este punto, la concentración desciende, denunciando la zona de la superficie donde el cobre había sido extraído. La concentración de plata (*línea gris*) también permanece relativamente constante hasta que el haz electrónico llega a unos cinco micrometros de la superficie (*b*). En ese punto, la plata desciende rápidamente, coincidiendo con el lavado de plata de la superficie. El oro, por el contrario, comenzó a subir de concentración a la misma distancia de la superficie en la que la concentración de cobre descendió. Continuó elevándose y alcanzó el máximo cuando el haz llegó al borde de la muestra (*c*).

cientes de cobre. Las superficies aparecen como si estuvieran formadas exclusivamente de oro. Sin embargo, cuando la plata forma parte también de la aleación, la superficie es rica en ambos metales preciosos; su color varía desde el blanco de la plata hasta el amarillo pálido. Si el orfebre quiere conseguir una superficie rica en oro tiene que extraer la plata.

Separar la plata del oro sin utilizar ácidos no es una operación sencilla, pero los moche lo pudieron hacer probablemente aplicando a la superficie de la lámina una mezcla de minerales corrosivos. De hecho, los experimentos realizados en nuestro laboratorio mostraron que la plata se desprende con bastante facilidad de la superficie de una chapa de aleación de oro, plata y cobre tratada con una pasta acuosa compuesta de sulfato férrico y sal. El sulfato férrico utilizado de este modo es casi tan efectivo como el ácido sulfúrico. El mismo procedimiento siguió siendo utilizado posteriormente por los orfebres de la cultura chimú, que dominaron la costa norte del Perú hasta que su reino fue conquistado por los incas hacia el año 1476 d. C.

De las grandes máscaras de momias andinas, impresionantemente doradas, que son familiares a los visitantes de los museos, muchas si no la mayoría son de manufactura chimú. El mayor ejemplar conocido, perteneciente en la actualidad a la colección del Museo Metropolitano de las Artes de Nueva York, está hecha de una lámina de metal y mide 74 centímetros de ancho por 40 de alto. Su superficie de oro es deslumbrante, pero nuestros análisis de laboratorio revelaron que el metal es una aleación de tumbaga que contiene solamente el 40 por ciento de oro: el resto está formado por un 48 por ciento de plata y 12 por ciento de cobre. Cuando fundimos una aleación con la misma composición, el lingote tenía un color amarillo pálido. Con todo, este objeto concreto es particularmente rico en oro. Algunos especialistas chimú consiguieron que parecieran de oro láminas de metal hechas con lingotes que contenían sólo un 12 por ciento de oro; el resto de la aleación era fundamentalmente cobre.



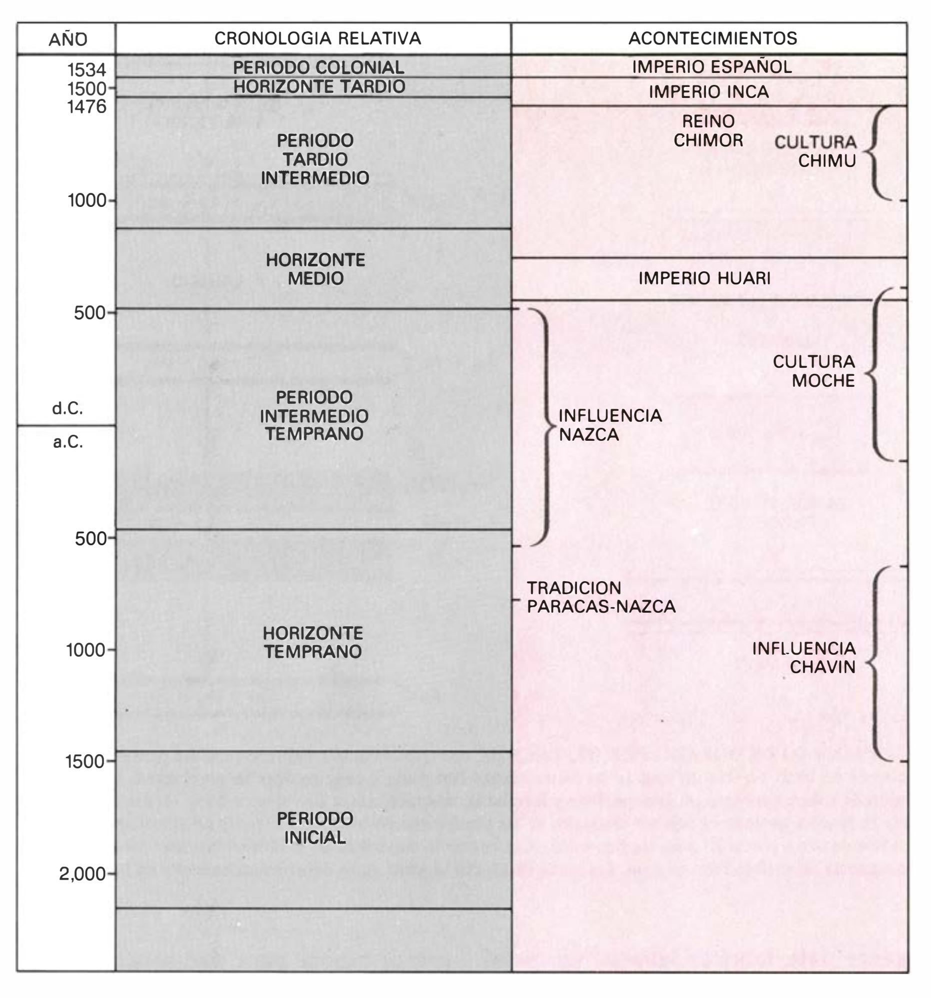
12. PROCESO DE DORADO POR DEPLECION, en representación esquemática. La primera etapa (*a*) consiste en batir en frío un lingote de cobre aleado con plata y oro, recocer luego el metal hasta que el óxido de cobre aparezca en la superficie y limpiarlo, después, antes de volver a batir de nuevo. A la vez que la lámina alcanza el espesor deseado, se ha ido formando una capa superficial compuesta por una mezcla de oro y plata. El paso siguiente (*b*) exige cubrir la superficie de la lámina con una pasta corrosiva compuesta de sulfato férrico y sal. La pasta disolverá la plata para dejar solamente oro en la superficie.

Las aleaciones de tumbaga, con sus propiedades inherentes para el enriquecimiento superficial de oro, eran comúnmente empleadas cuando los españoles se establecieron en México, Centroamérica y América del Sur, en el siglo XVI. Mientras que en los Andes centrales la tumbaga se usó principalmente para fabricar láminas de metal de aspecto dorado, en el norte de América del Sur y en América Central y México se fundía para conseguir formas tridimensionales.

Una vez que los objetos de tumbaga se funden, pueden ser dorados por depleción directamente disolviendo el cobre de la superficie con una solución corrosiva. No es necesario el recalentamiento. Pero si la aleación contiene poco oro son necesarios repetidos recalentados y lavados para obtener una capa superficial enriquecida consistente. A veces el orfebre yuxtaponía diferentes colores superficiales sobre el mismo objeto mediante el tratamiento químico de sólo algunas áreas seleccionadas de la tumbaga; de este modo se podía contrastar el dorado de la superficie tratada con el color más rojizo de la aleación no tratada. En algunas ocasiones los orfebres hicieron objetos con tumbagas que contenían diferentes proporciones de metales preciosos, fundiendo una aleación sobre otra para realzar el contraste de color de la superficie de diversas partes del objeto.

El desplazamiento electroquímico y el dorado o plateado por depleción no fueron los únicos métodos de tratamiento superfical de metales inventados en el Nuevo Mundo en la época precolombina para dar apariencia de metal precioso a un sustrato de metal poco noble. Los orfebres también doraban objetos con oro fundido y cubrían superficies de metal con panes de oro, pero recurrían a estos métodos en muy pocas ocasiones. De hecho, incluso el recubrimiento por desplazamiento electroquímico (una técnica esencialmente de "cobertura y ocultamiento") no tuvo más que un impacto local y poco duradero en las prácticas metalúrgicas precolombinas. Fue la metalurgia de depleción superficial (una técnica, por el contrario, de "desarrollo y engrandecimiento") la que dominó la tecnología metalúrgica del Nuevo Mundo durante casi dos milenios.

La explicación para esta elección de lo que sólo puede ser definido como un estilo tecnológico, se puede buscar en el mundo de las actitudes culturales y las ideologías. ¿Nos es posible reconstruir cuáles pueden haber sido las actitudes de los hombres de los Andes hacia los metales? En cierto modo sí. Consideremos el desarrollo de la metalurgia en el Viejo Mundo durante los tiempos prehistóricos. Es evidente que las importantes metalurgias de las edades del Bronce y el Hierro en Europa y el Cercano Oriente recibieron sus más poderosos estímulos desde los ámbitos



13. TABLA CRONOLOGICA que refleja las fases más importantes de la prehistoria andina, desde el año 2000 a. C. (Período Inicial) hasta el momento de la conquista española ocurrida el año 1534 d. C.

de la guerra, el transporte y la agricultura. Las armas y las herramientas fueron los principales objetos de la metalurgia del Viejo Mundo, y los especialistas metalúrgicos buscaron y desarrollaron las propiedades mecánicas que eran necesarias para la correcta ejecución de tales manufacturas: dureza, filo cortante, capacidad para mantener ese filo, resistencia al impacto y resistencia a la tensión.

En el Nuevo Mundo el énfasis puesto en la producción de metales era bastante distinto. Los objetos que constituían el centro de atención de la producción metalúrgica eran utilizados en los ámbitos de la vida asociados con el poder político, la ostentación del status social y la comunicación de las creencias religiosas. Tales objetos, los atributos de rango y poder tanto en el mundo de los vivos como en el de los muertos, incluían elaboradas orejeras (un tipo de pendientes), bezotes, objetos de culto y máscaras funerarias que se colocaban sobre la envoltura de las momias, tales como las que fabricaron los hombres de la cultura chimú. En términos antropológicos, en las sociedades del Nuevo Mundo los metales tenían importantes funciones simbólicas. El ímpetu de la metalurgia andina está en el desarrollo de las propiedades de los metales y aleaciones que contribuyeran a cumplir tales funciones. La propiedad más importante de todas era el color, y los dos colores fundamentales del espectro metalúrgico del Nuevo Mundo eran los de la plata y el oro.

Sabemos, por las crónicas que los españoles escribieron sobre la vida religiosa y política de los incas, que la familia real decía descender de la unión del sol y la luna. Los incas consideraban el oro como "el sudor del sol" y la plata como "las lágrimas de la luna". Ambos metales ocupaban, pues, un lugar preeminente en el origen mitológico de la dinastía real y estaban completamente controlados por el emperador. Los españoles se quedaron maravillados al ver las habitaciones de los palacios del Inca ocupadas con jardines interiores cuyas plantas, flores y pájaros eran del color de estos preciosos metales. La predilección andina hacia los colores de la plata y el oro no era una novedad introducida por los incas; mucho antes, hacia el año 1000 a. C., aparecen testimonios de ella con la difusión de los cultos religiosos de Chavín por amplias regiones de los Andes centrales. En el ritual de la cultura chavín se utilizaban profusamente objetos litúrgicos de oro; fue durante este período cuando se empleó por vez primera la aleación de cobre y plata.

Si la andina era una metalurgia orientada fundamentalmente hacia el color, también lo estaba hacia las superficies. El color de un objeto de metal está en su superficie. De hecho, tal como hemos visto, un objeto puede tener un color en la supeficie y otro bastante diferente bajo ella. No es extraordinario, por lo tanto, que los aspectos más innovadores de la metalurgia andina aparezcan como respuesta al afán de conseguir los colores del oro y la plata culturalmente apreciados. Tanto el desplazamiento electroquímico como el dorado o plateado por depleción daban los colores apropiados a objetos metálicos que no estaban hechos exclusivamente de alguno de los metales preciosos.

Dentro de este contexto, ¿cómo se puede explicar la persistencia de la metalurgia de depleción superficial como la técnica de coloración preferida en el Nuevo Mundo? ¿Por qué sobrevivió a todas las demás, difundiéndose ampliamente más allá de los Andes? Quizá porque era una técnica que hacía aparecer en la superficie de los objetos unas cualidades que ya estaban en su interior. El enriquecimiento de la superficie lo único que hacía era realzar sencillamente un aspecto del metal que le era propio; hablando en términos metalúrgicos, la aleación de que está hecho un objeto contiene el metal (oro o plata) que es posteriormente realzado o llevado a la superficie. Es el color dado a la superficie el que finalmente proporciona el carácter visual distintivo al objeto, comunicando mensajes culturales de poder, status y creencias religiosas. Pero, por otra parte, el color de la superficie no es otra cosa que la manifestación externa de una cualidad interna. Aunque nunca pueda darse una respuesta definitiva al problema, a pesar de que los factores económicos o tecnológicos contribuyeron al desarrollo de la metalurgia superficial en los Andes, fueron factores ideológicos los que más poderosamente influyeron en la dirección tecnológica elegida.

Extinciones masivas en los océanos

En el curso de breves intervalos de tiempo se han extinguido gran cantidad de animales y plantas marinas. Argumentos geológicos hacen pensar que la mayoría de esas extinciones masivas se debieron al enfriamiento del mar

Steven M. Stanley

A lo largo de la historia de la tierra las especies animales y vegetales se han ido extingiendo con un ritmo bastante regular. Sin embargo, en ocasiones se han dado valores anormalmente elevados de extinción específica. Esas extinciones masivas han afectado tanto a organismos continentales como marinos, aunque el registro fósil está mucho mejor documentado en lo que se refiere al medio marino. En el transcurso de intervalos geológicos breves, de sólo algunos millones de años, han desaparecido la mayoría de especies oceánicas y hasta la mitad de las familias. Un descalabro de tal magnitud sólo pueden inflingirlo cambios radicales del ambiente, a escala regional o, quizá, global. Para explicar esos hechos algunos autores han postulado, en los últimos años, causas extraterrestres, como radiaciones cósmicas derivadas de la explosión de una supernova. Otros han dirigido la atención a agentes más comunes: cambios drásticos de los "factores limitantes" del medio ambiente, entre los que cabe citar la temperatura y el espacio vital del fondo marino, que determinan de ordinario la distribución y abundancia de las especies en el mar.

El factor más importante a la hora de limitar la distribución geográfica de una especie animal en el mar es la temperatura del agua: una especie dada sólo puede sobrevivir dentro de un margen concreto de temperaturas. Los animales actuales constituyen una excelente confirmación de este hecho; una discontinuidad térmica marca con frecuencia el límite de la distribución geográfica de una especie. Basta un episodio de enfriamiento climático para acabar con cualquier especie que no esté adaptada a las nuevas temperaturas, más frías, o que carezca de un refugio más templado a donde emigrar.

En las páginas que siguen presentaré los nuevos argumentos que apuntan al enfriamiento climático como causa principal de la mayoría de crisis marinas conocidas. Son tres las pruebas fundamentales. Primera: ciertos indicadores independientes, así la presencia de gravas glaciares, manifiestan que la extinción masiva coincide en el tiempo con un enfriamiento (aunque no todos los enfriamientos tienen por qué estar asociados a glaciaciones). Segunda: el registro fósil sugiere que las especies adaptadas a aguas más cálidas o con una amplitud menor de tolerancia a la temperatura sufrieron el mayor deterioro, como cabe esperar de un descenso de la temperatura. Tercera: el registro fósil revela también que la mayoría de las extinciones marinas fueron graduales, no súbitas, prolongándose durante centenares de miles, y aún millones, de años. Este hecho cuadra con el ritmo que sigue un enfriamiento global, por lo general lento y episódico.

La más famosa de las extinciones masivas –la desaparición, a finales del Cretácico, hace unos 65 millones de años, de los dinosaurios, junto con la mayoría de las especies animales del mar– parece que fue, también, un proceso gradual. Como explicaré más adelante, éste constituye el argumento principal contra la atribución entera de la crisis cretácica al impacto de un asteroide, tesis esta defendida por Walter y Luis W. Alvarez y sus colegas, de la Universidad de Berkeley en California. Los efectos ambientales de un impacto son de corta duración, por lo que no pueden explicar por sí solos una crisis que se prolongó, por lo menos, dos millones de años. Además, aun aceptando que la hipótesis de los Alvarez fuera correcta, no hay razones para admitir que los impactos de asteroides justifiquen otras muchas crisis oceánicas.

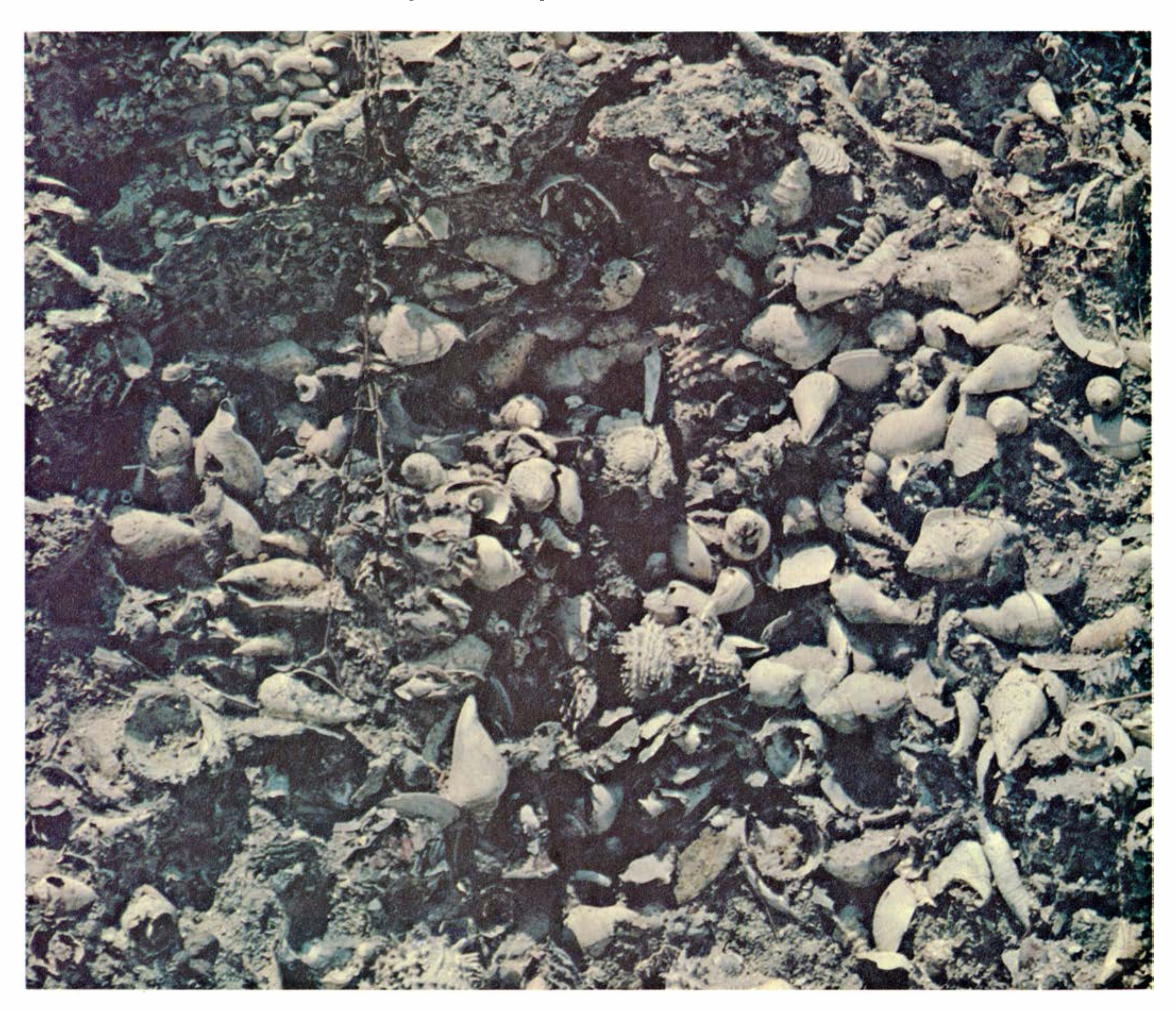
Antes de abordar las pruebas que ligan las extinciones en masa con el enfriamiento climático, permítaseme considerar los fallos de una explicación alternativa, que concede gran importancia a la alteración de otro factor limitante, a saber, el área del fondo marino somero, que decrece, cuando baja el nivel del mar, a causa de la emersión de buena parte de la plataforma continental. La hipótesis supone que para dar cabida a una población diversa de animales que vivan en el fondo y no estén adaptados a las condiciones de mares profundos se requieren grandes áreas de fondos marinos someros. No parece que sea así. Una plataforma continental bastante estrecha puede albergar una enorme diversidad de organismos de fondo. En este sentido, 3000 especies de moluscos con concha, representantes de centenares de familias, viven en los segmentos de la estrecha plataforma continental de la costa occidental del continente americano. Cerca de un millar de especies de moluscos conchíferos ocupan la pequeña área de fondo marino somero que bordea las islas Hawai.

Por otra parte, en estos últimos años ha aumentado notablemente el conocimiento de los cambios del nivel del mar en el pasado geológico; y nos dice que el "hacinamiento biótico" ocasionado por el descenso del nivel del mar no pudo haber causado la mayoría de las extinciones masivas conocidas. Gran parte de las pruebas proceden de la aplicación de técnicas de estratigrafía sísmica desarrolladas por investigadores de la Exxon Corporation. Reflejando ondas sónicas de gran potencia en las capas de sedimentos antiguos que se encuentran por debajo de las actuales plataformas continentales se detectan a menudo discontinuidades entre capas

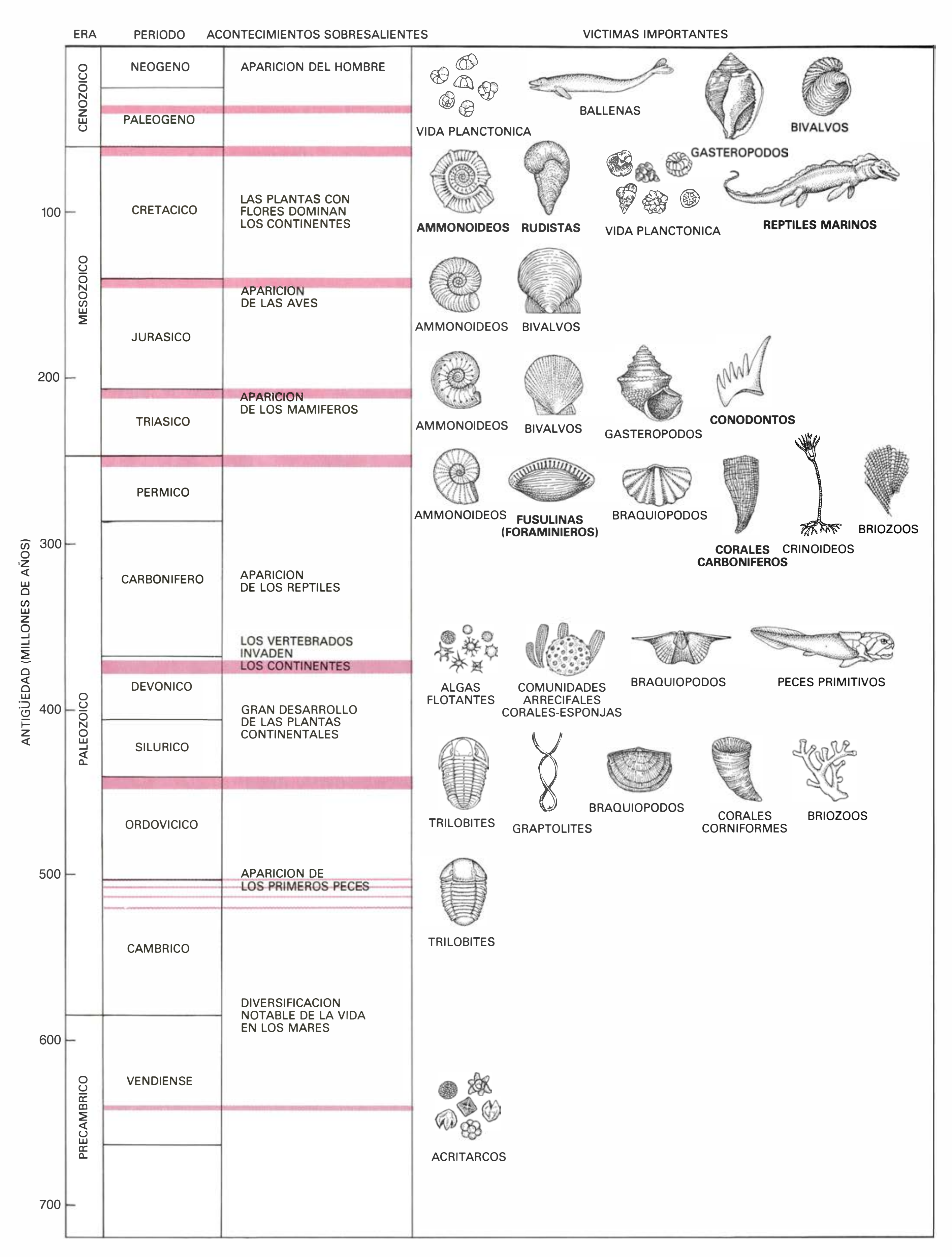
superpuestas, que difieren en densidad y, por consiguiente, refractan las ondas sónicas de modo diverso. (Una discontinuidad puede revelar el límite entre un lodo depositado en agua profunda y arenas aluviales que se acumularon posteriormente en una llanura costera, cuando el mar se había retirado a un nivel más bajo que el actual.) El análisis de rocas que afloran en el continente nos lleva a épocas en que el nivel del mar era más elevado.

Estos estudios han demostrado que durante muchas de las extinciones masivas el nivel del mar no era inferior al actual. Y a la inversa, con un nivel de mar bajo a menudo no se han dado crisis biológicas. Veamos un ejemplo. A finales del Eoceno, hace unos 40 millones de años, una crisis eliminó muchas especies marinas europeas precisamente cuando Europa se hallaba inundada de mares someros. A continuación, en el Oligoceno, los mares se secaron y el océano alcanzó el nivel más bajo, seguramente, de los últimos 100 millones de años. Aunque el nivel de las aguas permaneció bajo durante unos cinco millones de años, no hubo extinciones particularmente importantes en Europa ni en el resto del planeta. La disponibilidad de espacio vital en el fondo marino poco profundo quizá limite el número de especies que lo habiten, pero la relación entre espacio vital y número de especies es débil.

Más recientemente, el nivel del mar ha bajado hasta 100 metros en varias ocasiones en los últimos tres millones de años, cuando los casquetes glaciares en expansión mantenían enormes volúmenes de agua al margen del ciclo hidrológico. Gran número de especies desaparecieron de los océanos durante esos períodos glaciares, pero las pautas geográficas de las extinciones sugieren que en ello influyó más el enfriamiento que la disminución del fondo oceánico. La relación empieza a salir a la luz, pues durante mucho tiempo los paleontólogos no habían analizado con atención las extinciones masivas de la Edad del Hielo. Faltos de métodos precisos para datar las faunas marinas fósiles, habían supuesto que las faunas que contenían muchas especies extinguidas eran geológicamente más antiguas, y que las especies extinguidas habían desaparecido a una velocidad baja. El estudio de fósiles microscópicos, como los foraminíferos planctónicos (organismos flotantes parecidos a



1. MOLUSCOS FOSILES, víctimas de la Edad del Hielo, extraídos de una lumaquela de las arenas de Pinecrest, cerca de Sarasota, Florida. Durante los avances del hielo glacial que se sucedieron en los últimos tres millones de años, cerca del 70 por ciento de las especies de moluscos de la región se extinguieron. Entre los bivalvos, las especies adaptadas sólo a condiciones tropicales desaparecieron totalmente, mientras que algunas especies capaces de vivir en aguas templadas sobrevivieron. Estos hechos sugieren que fue el enfriamiento climático debido a la glaciación lo que causó las extinciones excesivas.

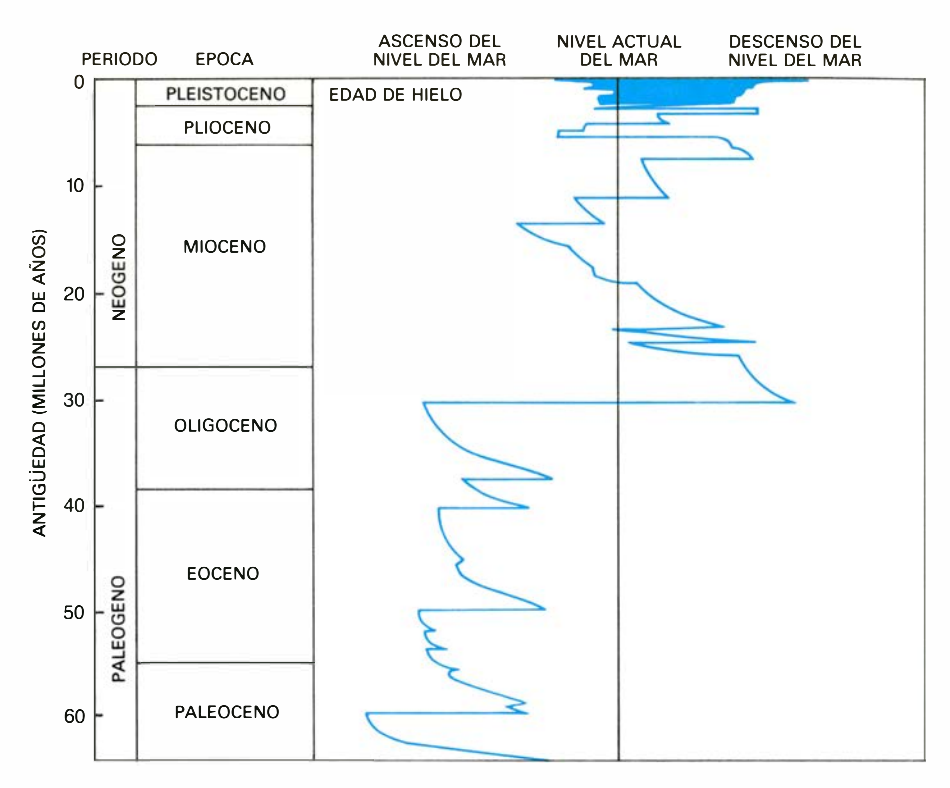


2. LAS EXTINCIONES MASIVAS (*líneas y barras de color*) **han afectado a todos los tipos de vida marina en los últimos 700 millones de años, desde algas unicelulares y plancton hasta enormes reptiles nadadores y ballenas. La crisis más famosa tuvo lugar hace unos 65 millones de años, a finales del Cretácico, y produjo la eliminación de la mayoría de especies marinas, a la vez que en el continente se extinguían los dinosaurios. En muchos casos, un grupo de animales o plantas logró recuperarse de una crisis y dar lugar, por evolución, a especies nuevas; en otros, el grupo entero desapareció del mar (*negrita*).**

las amebas, aunque dotados de caparazón), ha rendido mayor precisión en las determinaciones de edad, demostrando que las extinciones se agruparon en la Edad del Hielo. En muchas faunas fósiles de sólo cuatro millones de años de antigüedad se han extinguido más de la mitad de las especies.

Si las extinciones de la Edad del Hielo hubieran derivado de un descenso del nivel del mar, habrían afectado a las comunidades de aguas someras de todo el planeta. Sin embargo, quienes estudian los moluscos fósiles han encontrado una concentración de extinción superior a la media sólo en las regiones que bordean el Atlántico al norte del ecuador. Las faunas del Atlántico occidental y del Caribe sufrieron el golpe mayor, con pérdidas de alrededor del 70 por ciento de sus especies; en el mar Mediterráneo y en el mar del Norte se extinguieron aproximadamente el 30 por ciento de los moluscos. Por el contrario, no hay pruebas de extinción anormalmente alta durante los períodos glaciares en las faunas de moluscos de las costas pacíficas de California, Panamá y Japón.

Todo ello concuerda con la hipótesis del enfriamiento. Las especies que desaparecieron fueron las que estaban adaptadas a aguas cálidas, o las que tenían un margen estrecho de tolerancia térmica y no pudieron escapar de la extensión de enfriamiento asociada a la expansión glacial. Las faunas del Pacífico se salvaron; sencillamente, emigraban hacia el norte y hacia el sur al compás de las fluctuaciones de temperatura. Un vistazo al mapa permite comprobar que ni el Mediterráneo ni el Caribe ofrecen vías fáciles de escape. Durante las expansiones glaciales, al enfriarse estas cuencas tropicales, muchos moluscos se extinguieron, atrapados en mares gélidos. Las arenas de Pinecrest, en Florida central y meridional, son del Plioceno inferior (hace, aproximadamente, unos 3,5 millones de años) y nos ofrecen un ejemplo excelente de este proceso. Contienen una enorme fauna fósil de moluscos tropicales, entre los que se cuentan muchas especies de bivalvos. Las que han sobrevivido en la región son capaces de vivir, además de en los trópicos, también en aguas menos cálidas. Todas las especies estrictamente tropicales han desaparecido.

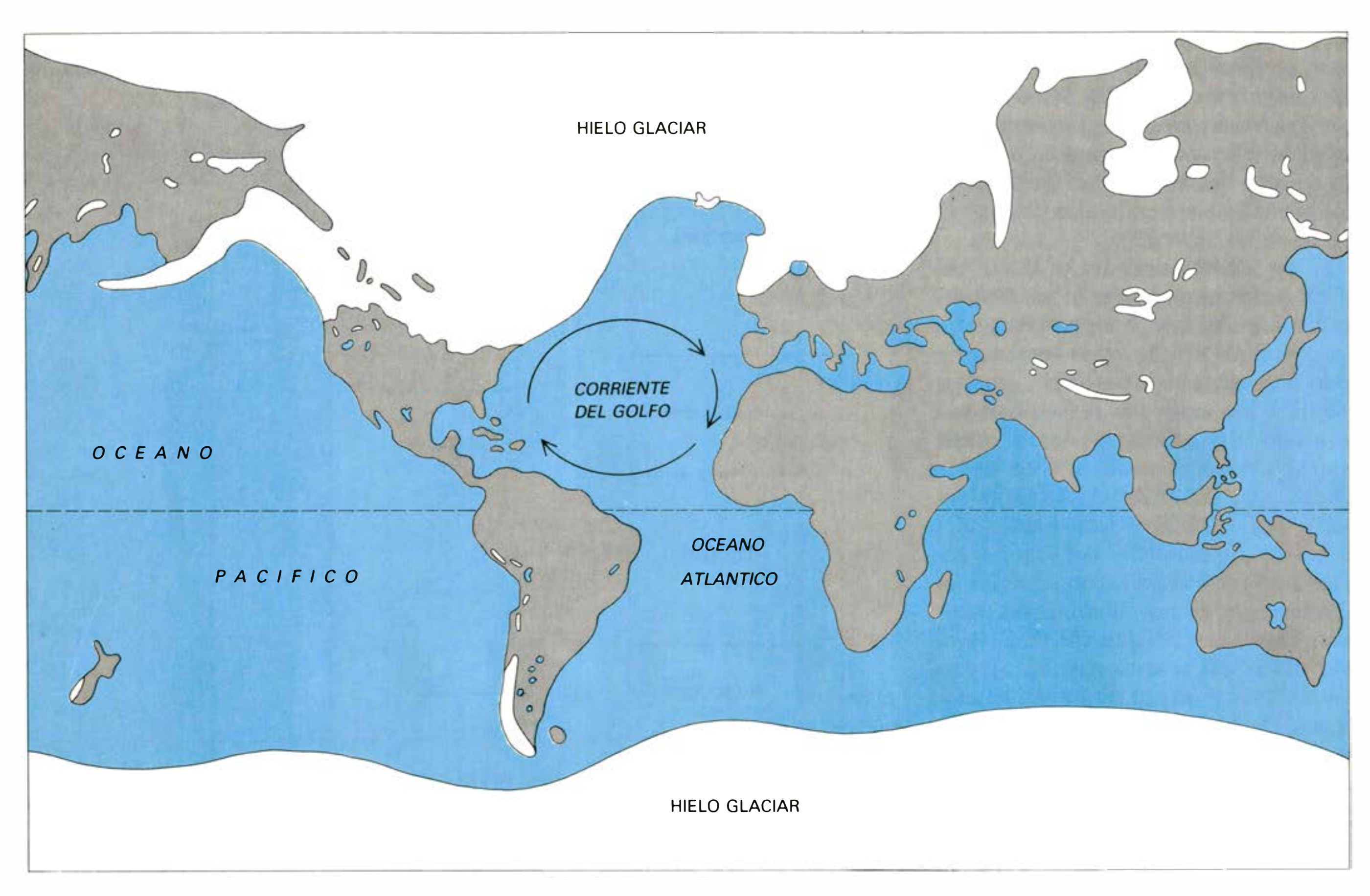


3. UN ESPECTACULAR DESCENSO DEL NIVEL DEL MAR, hace unos 30 millones de años, no produjo extinción masiva patente, en contradicción con la hipótesis que afirma que los descensos de las aguas provocan crisis marinas porque reducen el espacio vital en los fondos de aguas someras. Durante el Oligoceno, la superficie de los océanos llegó al nivel más bajo de los últimos 600 millones de años y permaneció por debajo del borde de la plataforma continental actual unos cinco millones de años. La ausencia de extinción anormalmente elevada en el Oligoceno contrasta con la crisis que eliminó muchas especies marinas en la época precedente, el Eoceno, cuando las aguas someras inundaban extensas áreas continentales. Los descensos del mar asociados a los avances de los hielos coincidieron con pulsaciones de extinción, aunque el modelo que siguen esas extinciones sugiere que las provocó el enfriamiento.

¿Hasta qué punto el enfriamiento climático da cuenta correctamente de las extinciones masivas anteriores? En la crisis del Eoceno superior citada anteriormente, que duró varios millones de años, las extinciones se extendieron más ampliamente que las de la Edad del Hielo. Además de la fauna europea, también sufrieron mucho los moluscos del Pacífico meridional y los de la costa del Pacífico noroccidental, que corresponde a los actuales Estados Unidos. Carole S. Hickman, de Berkeley, ha demostrado que en esta última área las especies pertenecientes a los grupos taxonómicos predominantemente meridionales (aguas cálidas) fueron las más perjudicadas. Richard Cifelli, de la Institución Smithsoniana, quien dio cuenta de la crisis en primera instancia, observó que había eliminado la mayoría de los tipos de foraminíferos planctónicos, pero que el grupo de foraminíferos de las globigerinas sobrevivió y llegó a predominar en el Oligoceno. Hoy las globigerinas son el único grupo de foraminíferos plactónicos que se encuentran ordinariamente en aguas frías: las especies que medran en los mares cálidos han evolucionado en su gran mayoría a partir del Eoceno. El florecimiento de las globigerinas, coincidente con la desaparición de las formas análogas de aguas cálidas, tiene explicación si se acepta que la crisis fue provocada por enfriamiento.

Hay muchas razones a favor de la tesis que afirma que las extinciones estuvieron acompañadas de un descenso de temperaturas. Las hojas fósiles encontradas cerca de Puget Sound documentan el cambio de bosques lluviosos casi tropicales a bosques templados por la misma época en que los moluscos desaparecían de las aguas próximas a la costa. Analizando el cambio de la proporción de especies de plantas con hojas de bordes lisos respecto de las de bordes dentados en una región determinada puede estimarse la variación de la temperatura media; las hojas de bordes lisos son más frecuentes en los climas tropicales, mientras que las hojas dentadas predominan en las zonas templadas. Aplicando ese método, Jack Wolfe, del Servicio Geológico de los Estados Unidos, ha calculado que la temperatura media anual en el Pacífico noroccidental descendió unos 10 grados Celsius durante el Eoceno superior. Un análisis parecido ha demostrado un enfriamiento semejante en Gran Bretaña.

Como en la Edad del Hielo, el enfriamiento eocénico parece que se originó en los polos y se extendió hacia el ecuador de forma irregular. Los fósiles encontrados en sedimentos profundos



4. EXTINCION EN MASA DE LA EDAD DEL HIELO. Quedó limitada a las regiones que bordean el Atlántico Norte. Cerca del 70 por ciento de las especies de moluscos del Atlántico occidental y del Caribe, y un 30 por ciento de las del Mediterráneo y del mar del Norte, desaparecieron. Durante los episodios fríos, las faunas de moluscos quedaron atrapadas en esas cuencas y no hallaron una vía de escape fácil hacia mares más cálidos. Por el contrario, los moluscos del Pacífico pudieron emigrar hacia el sur cuando se produjo el descenso de la temperatura, con lo que se salvaron de una extinción importante. La trayectoria de la Corriente del Golfo, muy al sur de la posición actual, permite ver la extensión del enfriamiento del Atlántico hace 18.000 años.

cerca de la Antártida revelan que los glaciares se expandieron ya avanzada esa época. Se produjo entonces un cambio drástico en la estructura térmica de los océanos, como han demostrado Richard H. Benson, de la Institución Smithsoniana, y otros investigadores. Tan pronto como el agua que bordeaba la Antártida se enfrió, parte de ella se hundió, dada su mayor densidad, y avanzó hacia el norte, formando una capa fría en el fondo del mar, la psicrosfera, que ha persistido hasta hoy. La separación de Groenlandia de Europa, durante el Eoceno, pudo ayudar también a la creación de la psicrosfera, permitiendo el desplazamiento de aguas gélidas del Artico hacia el Atlántico.

No hay razón para suponer que los importantes enfriamientos que ocurrieron en épocas anteriores partieron, necesariamente, de los polos. Más bien parece que tuvieron carácter global, con un descenso significativo de temperaturas tanto en el ecuador como en latitudes elevadas. Tal enfriamiento universal debió perjudicar principalmente a las biotas tropicales. Mientras que las biotas de latitudes altas podrían emigrar hacia el ecuador, mientras no se interpusieran barreras geográficas, siguiendo la zona climática a la que estaban adaptadas, no habría refugio más cálido para las especies tropicales, pues habitaban ya las áreas más caldeadas. Además, las especies tropicales son en muchos casos menos tolerantes a los cambios térmicos que las especies de regiones templadas, donde, en general, son más fuertes las fluctuaciones estacionales. Un enfriamiento que produjera un descenso por debajo de los 18 grados Celsius eliminaría muchas especies marinas tropicales del mundo actual. De haber contribuido un enfriamiento global a las crisis marinas anteriores a la era cenozoica, era de los mamíferos, que empezó hace 65 millones de años, las extinciones masivas deben reflejar una desviación en la abundancia de organismos tropicales. El examen del registro fósil en ese sentido permite constatar que así es.

Los mejores datos en favor de esa tesis proceden de la extinción en masa que tuvo lugar al término de la era mesozoica, o era de los dinosaurios, y al inicio de la cenozoica. El último período del Mesozoico fue el Cretácico. Durante la parte superior del Cretácico, mares de poca profundidad cubrían grandes áreas continentales hoy emergidas; una franja de mares tropicales, conocida por Tethys, se extendía del Asia sudoriental al golfo de México, atravesando la región mediterránea. Entre los organismos marinos, los más perjudicados por la crisis del Cretácico terminal fueron las faunas tethyanas.

Se diezmaron muchos grupos de diminutas algas flotantes, entre los que se cuentan la mayoría de cocolitofóridos (organismos esféricos unicelulares que segregan unas estructuras con aspecto de escudo, de carbonato cálcico, que se desprenden cuando los organismos mueren). Antes de la crisis, el carbonato cálcico precipitaba en gran cantidad hacia el fondo oceánico, dando lugar a potentes depósitos de creta (los acantilados blancos de Dover son ejemplo de ello), de los cuales deriva el nombre de Cretácico. Los cocolitofóridos no han vuelto a recobrar su eminente posición en los mares tropicales. Los foraminíferos planctónicos también sufrieron pérdidas notables: sus únicos supervivientes fueron las globigerinas, el mismo grupo que llegó a predominar después de la crisis del Eoceno y que hoy prospera en los mares fríos.

La desaparición, en el Tethys, de los rudistas, bivalvos constructores de arrecifes que se parecían a corales con conchas cónicas, fue particularmente espectacular. Tanto prosperaron esos animales en el Cretácico que, al parecer, empujaron a los corales a un papel subordinado en los arrecifes tropicales. De no ser por la súbita extinción de los rudistas, sin duda dominarían éstos, y no los corales, los arrecifes que puntean los mares someros tropicales del mundo actual. En el resto del fondo marino del Tethys desaparecieron otros grupos de bivalvos y gasterópodos (caracoles), así como familias de macroforaminíferos bentónicos. Más desafortunada, en razón de su gran belleza, fue la extinción de los ammonoideos: moluscos cefalópodos nadadores cuyas conchas espiraladas divididas en cámaras, como los de las formas próximas actuales, el nacarado nautilo, albergaban un cuerpo parecido al calamar. Los ammonoideos representan un grupo cuya eliminación no refleja la tendencia tropical general de la crisis terminal del Cretácico: desaparecieron tanto las especies no tropicales como las tropicales.

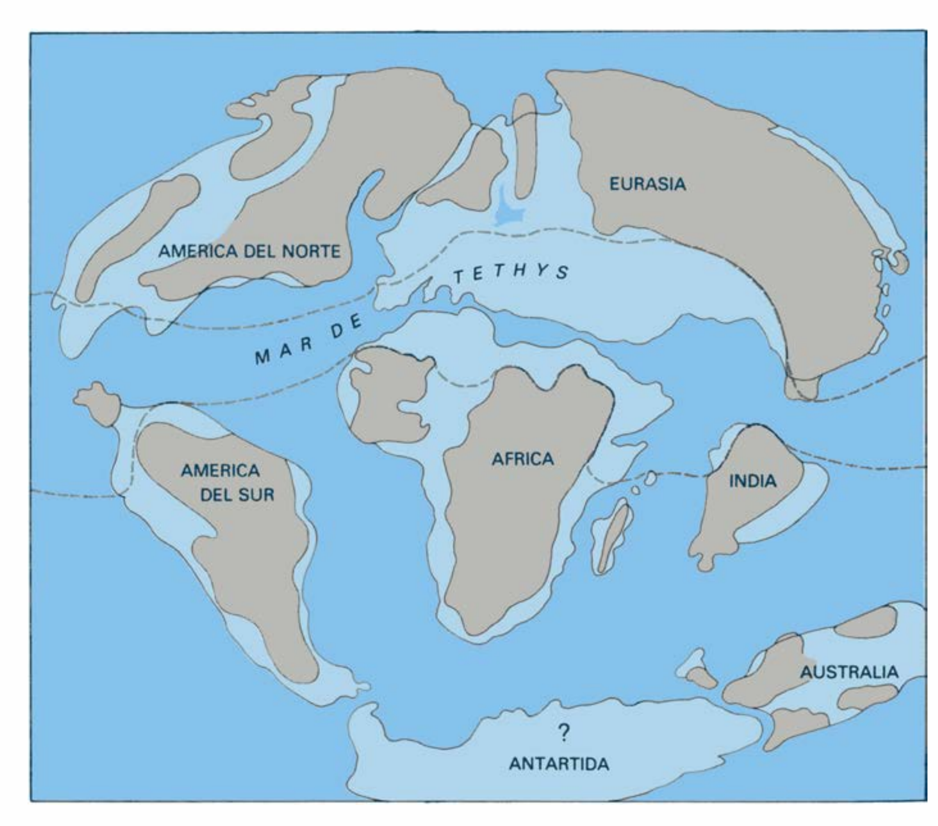
No obstante, la desviación es real, como lo demuestra la confusión existente acerca de la datación correcta de ciertas formaciones rocosas de latitudes elevadas. Durante muchos años, algunos paleontólogos asignaron rocas danesas del Daniense al Cretácico, porque contenían grupos fósiles del Cretácico. Muchos investigadores suponen hoy que la crisis cretácica terminal precedió al Daniense. Sin embargo, su efecto fue en Dinamarca tan pequeño que muchas especies cretácicas sobrevivieron hasta el Paleogeno, primer período de la era Cenozoica. (El Cenozoico se ha dividido tradicionalmente en período terciario, que se inicia hace 65 millones de años, y cuaternario, que comprende los últimos dos millones de años. Sin embargo, muchos paleontólogos preferirían dividir el Cenozoico en dos partes más igualadas: el período paleogeno, que llegaría hasta el final del Oligoceno, hace 26 millones de años, y el período neogeno.) De manera parecida, la formación Cannonball de Dakota del Norte, atribuida en principio al Cretácico, se ha demostrado de edad paleogena. Aunque la crisis cretácica destruyó la mayor parte de las especies de bivalvos de la región del golfo de México, cerca del 60 por ciento de las especies de bivalvos que se encuentran en la formación Cannonball se han hallado también en depósitos cretácicos varios millones de años más antiguos. Esa es más o menos la tasa de supervivencia en intervalos en los que no ha habido crisis.

No sólo pagaron las biotas tethyanas un precio mucho mayor que las no tropicales; además, las biotas de latitudes nórdicas frías se desplazaron hacia el sur, ocupando los lugares de la flora y fauna tropicales desaparecidas. Por ejemplo, los caracoles tethyanos quedaron diezmados con la extinción masiva; Heinz Kollman, del Museo de Historia Natural de Viena, ha encontrado que, durante el Paleogeno inferior, los caracoles, en una amplia región que va de Groenlandia a Africa septentrional, pertenecían a grupos que durante el Cretácico habían vivido sólo en los mares nórdicos.

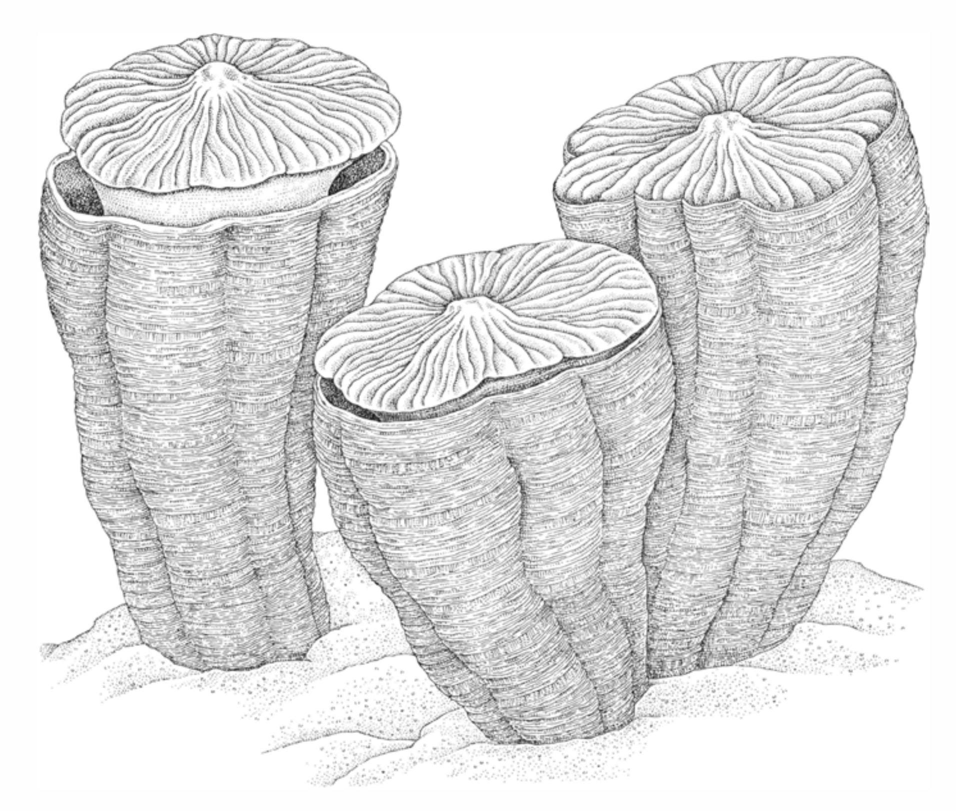
Leo J. Hickey, de la Universidad de Yale, ha demostrado hace poco, mediante el análisis de los bordes de hojas fósiles, que la región de Wyoming experimentó una pulsación de enfriamiento durante la transición del Maastrichtiense (última época del Cretácico) al Paleoceno. Se sabe que el Maastrichtiense es el intervalo más frío del Cretácico, pero en Wyoming las floras del Paleoceno inferior dan fe de un clima que fue hasta 10 grados Celsius más frío aún. Parece que las bajas temperaturas persistieron en Wyoming varios millones de años.

Dejamos aquí las pruebas que permiten relacionar las crisis cretácica con un enfriamiento climático. En los últimos años ha atraído la atención del gran público la hipótesis de Alvarez según la cual el impacto de un asteroide de unos 10 kilómetros de diámetro exterminó organismos marinos y dinosaurios. Se supone que el impacto levantó una nube de polvo que rápidamente envolvió la tierra, precipitando luego sobre el suelo desde la atmósfera. El principal argumento de esta suposición radica en el descubrimiento, en las sucesiones estratigráficas de todo el mundo, de una concentración significativa de iridio en una delgada capa sedimentaria del límite Cretácico-Paleogeno. El iridio, metal del grupo del platino, es raro en la tierra, pero relativamente abundante en los meteoritos. Se acepta que la mayor parte de meteoritos son fragmentos de asteroides.

¿Cómo pudo un impacto haber causado extinciones en el océano? Se han propuesto diversos mecanismos, entre los que cabe citar el del aumento brusco, o el descenso igualmente espectacular, de la temperatura. Los Alvarez han sugerido que la nube de polvo pudo ocasionar un período de tres a seis meses de oscuridad; la detención de la fotosíntesis sacrificaría el fitoplancton oceánico. Los efectos de esa extinción habrían producido una reacción subsiguiente en la cadena alimentaria, elimi-



5. UNA FRAJA OCEANICA TROPICAL, el mar de Tethys, cubría grandes extensiones de los actuales continentes en el Cretácico superior. Las especies tethyanas sufrieron la extinción mayor en la crisis de finales del Cretácico. Que la extinción incidiera especialmente en el área tropical permite suponer que la causó un enfriamiento. La línea a trazos indica los límites entre condiciones templadas y tropicales.



6. RUDISTAS, moluscos bivalvos que, en el Cretácico, desplazaron a los corales del papel de constructores arrecifales predominantes en el mar de Tethys; desaparecieron definitivamente al final de ese período. Su declive comenzó unos dos millones de años antes de su extinción, de lo que se deduce que ésta no fue consecuencia de un acontecimiento súbito, como podría ser el impacto de un asteroide. Al revés de lo que acontece con los corales, los rudistas sólo estaban unidos por sus conchas, no por sus tejidos blandos. Aunque no solían sobrepasar los 20 cm de altura, algunos alcanzaban alturas superiores a un metro.

nando organismos marinos de tamaño mayor.

La hipótesis de Alvarez contiene una predicción verificable: si el impacto del asteroide causó extinciones al final del Cretácico, éstas debieron haber sido súbitas, es decir, prácticamente instantáneas en la escala del tiempo geológico. Pero, a juzgar por los datos recientes, no parece ser ese el caso. Los depósitos del Cretácico terminal de Stevns Klint, en Dinamarca, se han tomado por demostración palmaria de la existencia de extinciones rápidas. Sin embargo, la parte más alta de estos depósitos es un "suelo duro" (*hard-ground*): un área de fondo marino en la que los sedimentos que permanecieron sin enterrar durante largo tiempo, a causa de la no deposición de nuevos sedimentos, se cementaron en una costra dura por precipitación de carbonato cálcico. Ahora bien, si no hay deposición de sedimento no hay registro fósil. Por consiguiente, el suelo duro de Stevns Klint hace pensar que los depósitos que haya en él no documentan realmente las últimas decenas de miles de años del Cretácico. Si ello es así, no apoyan la afirmación según la cual las extinciones del Cretácico terminal se extendieran en el curso de sólo algunos milenios, como máximo.

En Zumaya, en la costa norte española, existe una sucesión aparentemente más continua. Depósitos de mar profundo, izados hasta la superficie por el movimiento de las placas, atraviesan el límite Cretácico-Paleogeno sin que nada haga pensar en alguna interrupción de la deposición sedimentaria; parece, pues, que puedan suministrar una historia más precisa de algunas extinciones cretácicas importantes. Las extinciones que parecen compatibles con la hipótesis del asteroide son las de los cocolitofóridos y de plancton del mismo tipo. De acuerdo con Stephen F. Percival, Jr., y Alfred G. Fischer, entonces en la Univeridad de Princeton, las especies cretácicas típicas de ese plancton desaparecieron súbitamente, por lo menos como formas abundantes, en un intervalo de menos de 10.000 años inmediatamente antes de que se formase la capa rica en iridio.

Por el contrario, Peter Ward, de la Universidad de California en Davis, y Jost Wiedmann, de la Universidad de Tübingen, han comprobado que la fauna de Zumaya sufrió una decadencia gradual en la última parte del Cretácico, y no una extinción súbita a finales de ese período. El número de especies de ammonoideos descendió de 10 a ninguna en un intervalo de unos dos millones de años, representados por unos 120 metros de sedimento. El ammonoideo más moderno se ha encontrado unos 12 metros por debajo de la capa de iridio, lo que hace suponer que la última especie se extinguió unos 100.000 años antes del supuesto impacto del asteroide.

Ward y Wiedmann han encontrado un decrecimiento semejante de los inocerámidos, grupo de bivalvos que desaparecieron cerca del final del Cretácico; el último inocerámido que han encontrado en Zumaya está 60 metros por debajo de la capa de iridio, lo que viene a corresponder a un millón de años más de antigüedad. Estos resultados confirman los del estudio global de los inocerámidos realizado por Annie V. Dhondt, del Real Instituto de Ciencias Naturales de Bélgica, que permite afirmar que tales bivalvos menguaron durante el Maastrichtiense. De manera parecida, según Erle G. Kauffman, de la Universidad de Colorado en Boulder, los rudistas (que edifican arrecifes en mares tropicales y no aparecen, por tanto, en los depósitos de mar profundo de Zumaya) quedaron diezmados en el Maastrichtiense medio, quizá dos millones de años antes de que se formara la capa de iridio. En las rocas del Maastrichtiense superior, la comunidad fósil de los rudistas se ofrece muy empobrecida, y no hay arrecifes importantes.

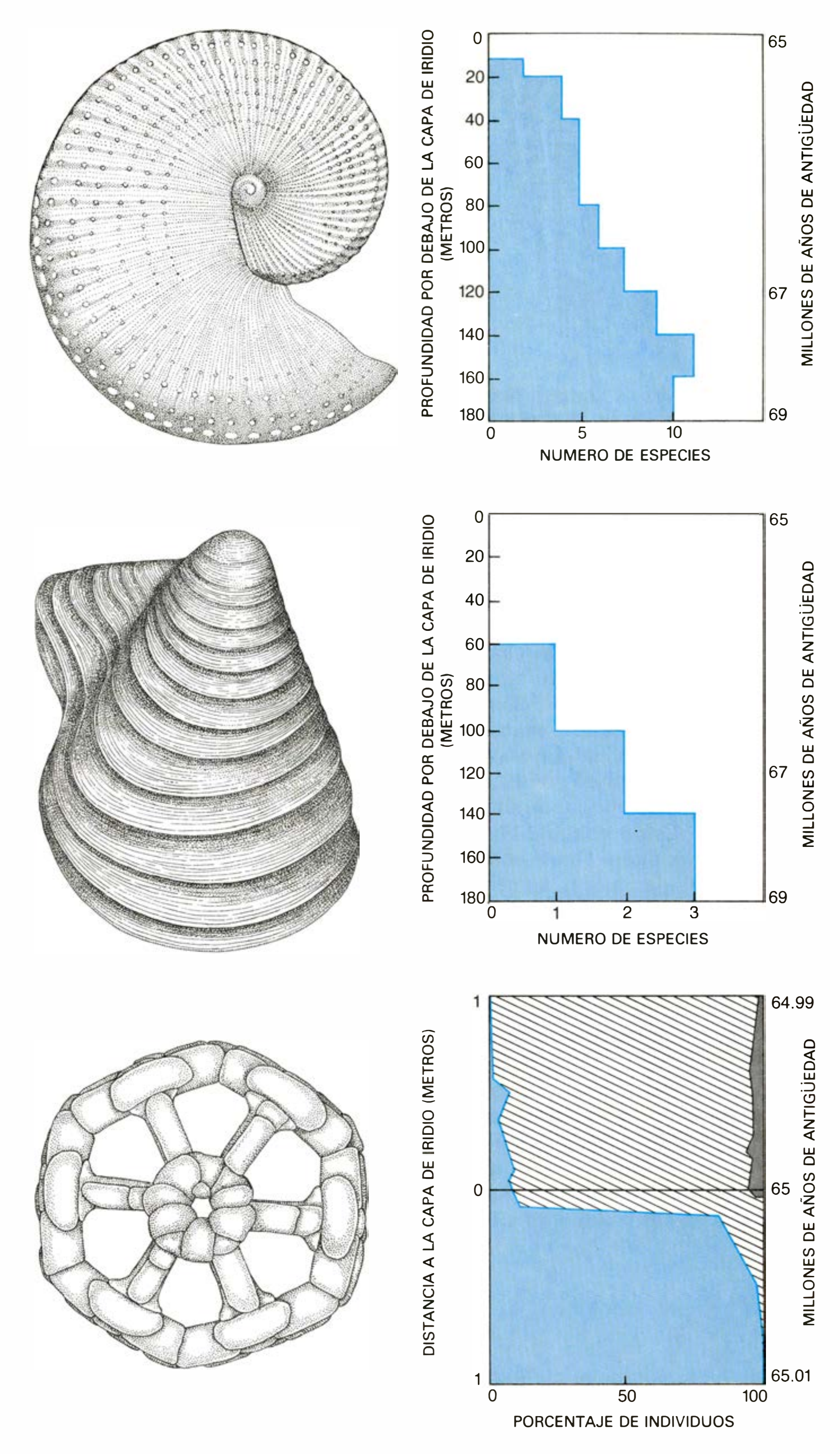
Resumiendo, la crisis cretácica no fue un acontecimiento breve. Diversos grupos de organismos entraron en un proceso de decadencia y se extinguieron, en momentos diferentes, a lo largo de un período de dos millones de años, por lo menos. La secuencia de las desapariciones es también significativa, ya que contradice la vieja idea, adoptada por los Alvarez, de que las extinciones empezaron en la parte baja de la cadena alimentaria y se propagaron hacia arriba en una suerte de caída de fichas de dominó. El humilde plancton resultó afectado muy al final de la crisis cretácica, después de la decadencia de muchos grupos de moluscos comedores de plancton y después de la total desaparición de los ammonoideos carnívoros.

Recientemente, los paleontólogos han propuesto algunas modificaciones a la hipótesis de Alvarez que permitirían dar cuenta del hecho de que las extinciones del Cretácico no fueron instantáneas. Esas nuevas hipótesis, que por ahora carecen de datos sólidos, se

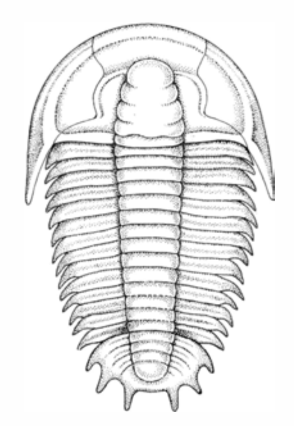
han establecido para explicar la afirmación de David M. Raup y J. J. Sepkoski, Jr., de la Universidad de Chicago, según la cual, y basándose en el análisis estadístico de la desaparición de más de 500 familias marinas, las extinciones de los últimos 250 millones de años han seguido ciclos de 26 millones de años. De acuerdo con una de las nuevas hipótesis, una pequeña y fría estrella acompañante daría vueltas alrededor del sol en una órbita muy alargada. Si ese cuerpo pasara cerca del sol cada 26 millones de años, podría perturbar las órbitas de los cometas y, durante un período amplio de tiempo, porciones de algún cometa podrían chocar contra la tierra, provocando las extinciones masivas. Algunos investigadores están empeñados en la búsqueda de la estrella acompañante; otros examinan los cráteres terrestres verosímilmente causados por impactos de cometa para comprobar si las edades de dichos cráteres siguen el mismo período que las extinciones masivas marinas.

El supuesto ciclo de 26 millones de años requiere también una investigación cuidadosa. En primer lugar, el último máximo de extinción, de hace unos 13 millones de años, es muy débil, y seguramente no lo habríamos detectado si hubiera sido más antiguo; quizás anteriormente se dieran muchos otros acontecimientos de ese tipo que hayan oscurecido el registro fósil y que no se ajusten al ciclo. Además, los máximos anteriores a los 90 millones de años que identifican Raup y Sepkoski no encajan bien en el ciclo. Puede que los máximos no sean realmente periódicos, sino que, simplemente, se sitúen más cerca de un ciclo de 26 millones de años que de una distribución al azar.

Tal hecho no requiere apelar a ninguna explicación extraterrestre. Podría deberse a que, tras cada extinción masiva, transcurriera cierto tiempo antes de presentarse otra crisis. El cambio ambiental que precipitó la crisis primera pudo durar algunos millones de años, obstaculizando la recuperación biótica que debe preceder a una nueva pulsación de extinción. Aun cuando mejorase el medio ambiente, los biotas diezmados necesitaron millones de años para volverse a diversificar hasta el punto de originar especies particularmente vulnerables a una nueva extinción en masa. Después de la crisis cretácica, por ejemplo, y después de la extinción masiva de amplitud parecida que se dio en el Pérmico superior, los biotas marinos quedaron empobrecidos durante varios millones de años.



7. PAUTAS DE EXTINCION en una sucesión estratigráfica de Zumaya, Guipúzcoa. Aportan pruebas en favor y en contra de la hipótesis de que el impacto de un asteroide provocó la crisis del Cretácico terminal. La prueba principal de la existencia del impacto es una capa fina de arcilla del límite Cretácico-Paleogeno en la que se aprecia una abundancia anormal del metal raro iridio. Los grupos extinguidos por el impacto de un asteroide deberían desaparecer súbitamente cerca de la capa de iridio. De acuerdo con el estudio de la formación de Zumaya, realizado por Peter Ward, de la Universidad de California, tanto los ammonoideos (*arriba*) como los inocerámidos, moluscos bivalvos (*centro*), presentan un empobrecimiento gradual en número de especies durante el Cretácico superior; ambos grupos desaparecieron antes de la capa de iridio y, por consiguiente, antes del supuesto impacto. Por el contrario, la extinción de algunas especies de cocolitofóridos (*abajo*) y de plancton calcáreo se produjo de manera brusca. De acuerdo con un cómputo de individuos efectuado por Stephen F. Percival, Jr., y Alfred G. Fischer, las especies cretácicas típicas (*color*) fueron remplazadas por especies oportunistas (*rayado*) y especies de nueva aparición (*gris*) en menos de 10.000 años, más o menos contemporáneamente a la deposición del iridio.



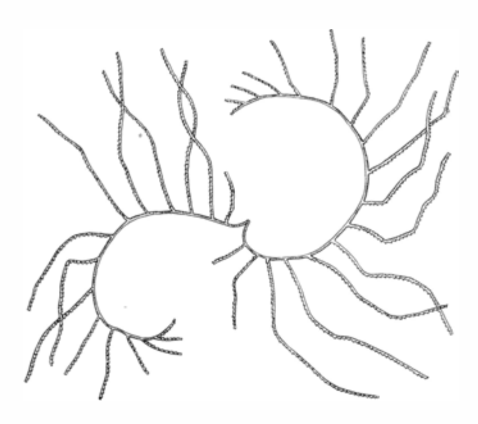
8. LAS FAUNAS DE TRILOBITES de mares someros tropicales se diezmaron repetidamente en extinciones masivas del Cámbrico, período que empezó hace 570 millones de años. Sobrevivieron especies adaptadas a aguas profundas y más frías.

Con independencia de si las extinciones siguen pautas cíclicas, existe una serie amplia de pruebas que relacionan las crisis, incluidas las anteriores al Cretácico, con el enfriamiento climático. El primer acontecimiento conocido tuvo lugar hace unos 650 millones de años, en el Precámbrico, cuando aún era escasa la vida animal. En esa época, de acuerdo con Gonzalo Vidal, de la Universidad de Lund, y Andrew H. Knoll, de la Universidad de Harvard, las extinciones diezmaron la población oceánica de los acritarcos, fitoplancton unicelular que parece haber sido el primer grupo de organismos que tuvieron células con núcleo [véase "Microorganismos planctónicos fósiles", por Gonzalo Vidal; INVESTIGACIÓN Y CIENCIA, agosto de 1983]. La desaparición masiva de los acritarcos, que más tarde se recuperaron hasta llegar a ser el grupo más importante de fitoplancton fósil conservado en las rocas del Paleozoico, coincide con un período en el que los glaciares recubrían muchas áreas de la tierra. Pruebas de la actividad glaciar lo son las formaciones arenosas que contienen grandes bloques que sólo pueden haber sido depositados por glaciares y, también, las estrías dejadas por los bloques sobre el suelo rocoso más antiguo al desplazarse estos bloques por el movimiento del hielo. Tales pruebas se han encontrado en todos los continentes, excepto en la Antártida (cuya superficie está en gran parte recubierta por hielo), de lo que se deduce que el enfriamiento climático del Precámbrico fue un fenómeno general.

En el período siguiente, el Cámbrico, abundaron en los mares animales invertebrados con concha. Los primeros invertebrados predominantes fueron artrópodos, los trilobites, así llamados por tener sus cuerpos segmentados divididos en tres lóbulos, uno central y dos laterales. Los trilobites aparecieron aproximadamente hace 530 millones de años y sufrieron una serie de extinciones masivas, de las cuales la mejor documentada se refiere a faunas que ocupaban los fondos marinos tropicales de Norteamérica. Allison R. Palmer, de la Sociedad Geológica de América, ha encontrado pruebas de tres crisis distintas, separadas por intervalos que van desde 5 hasta 7,5 millones de años. Por la potencia de los sedimentos depositados durante uno de esos acontecimientos ha concluido que se extendió a lo sumo de 4000 a 6000 años. Parece, por tanto, que a las breves pulsaciones de extinción siguieron largos intervalos durante los cuales los trilobites volvían a expansionarse, originando muchas especies nuevas antes de que llegara el golpe de una nueva crisis. Cada vez la nueva fauna evolucionaba a partir del mismo grupo ancestral, los trilobites olénidos, cuyos fósiles se encuentran en sedimentos de aguas profundas de mar adentro, presumiblemente frías. Este modelo repetido se explica fácilmente por el enfriamiento climático; cuando los breves períodos de mayor frío diezmaban periódicamente las faunas tropicales de aguas poco profundas, los olénidos adaptados a temperaturas frías "esperaban su entrada en escena", mar adentro, de donde partían para invadir los mares someros vacíos y originaban nuevas especies de trilobites que se adaptaban a las condiciones tropicales restablecidas.

Los trilobites fueron también víctimas de una extinción masiva de finales del Ordovícico, hace unos 440 millones de años. La crisis ordovícica eliminó unas 100 familias de animales marinos, sobre todo, de nuevo, en los trópicos. Su impacto sobre los graptolites es particularmente instructivo.

Los graptolites fueron un grupo raro. Vivían en colonias flotantes de individuos, unidos entre sí por un tallo común. Seguramente comían plancton y otras partículas alimenticias pequeñas. Los sedimentos fósiles que se encuentran a lo largo de los bordes del Atlántico Norte revelan que, varios millones de años antes del fin del Ordovícico, distintas especies de graptolites medraban en varias franjas ecológicas paralelas al ecuador. Al final del Ordovícico, cuando se produjeron las extinciones, las franjas se habían reducido a una única provincia biogeográfica de la región ecuatorial. Simultáneamente, la actividad glacial, que se había iniciado poco antes, conocía su apogeo; el hielo se extendía en varias direcciones a partir de Africa septentrional. La redistribución geográfica de los graptolites sugiere que sus extinciones guardaban relación con el enfriamiento de su hábitat.



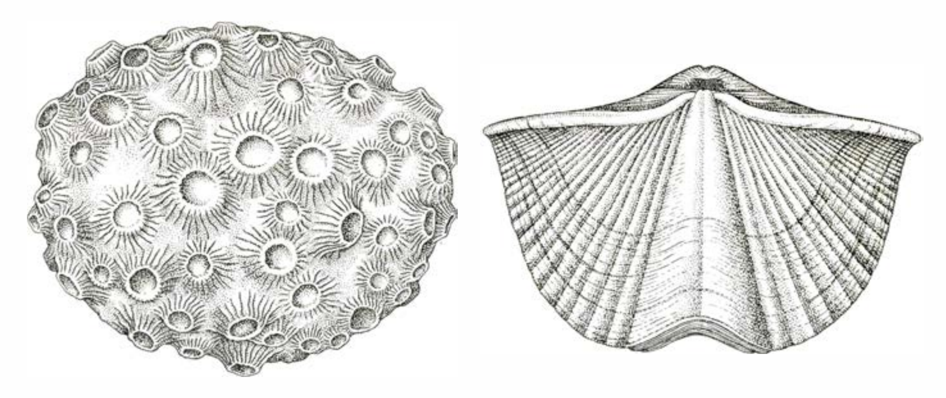
9. GRAPTOLITES, animales coloniales flotantes que sufrieron una fuerte disminución a finales del Ordovícico, cuando los glaciares avanzaban desde un área central localizada en el norte de Africa. Cada "diente" del tallo albergaba un individuo.

Las especies tropicales sufrieron otro duro golpe en una crisis marina de finales del Devónico, hace unos 370 millones de años. Los corales primitivos y las esponjas constructoras de arrecifes calcáreos que se habían desarrollado anteriormente en el mismo período sufrieron una pulsación de extinción de la cual jamás se recuperarían totalmente; muchos otros grupos marinos tropicales desaparecieron al mismo tiempo. Por el contrario, los grupos animales que vivían cerca del polo Sur se resistieron poco, como han demostrado las investigaciones de Paul Copper, de la Laurentian University de Ontario. América del Sur estaba junto al polo sur durante el Devónico, y su extraña fauna, que carecía de constructores de arrecifes y de especies de aguas cálidas, sobrevivió con pocas pérdidas. Los braquiópodos (bivalvos que se encuentran tanto en aguas tropicales como templadas) resultaron muy perjudicados por la crisis, pero no de una manera indiscriminada. El 90 por ciento de las familias de braquiópodos que se extinguieron carecían de representantes en las aguas frías que cubrían amplias áreas de América del Sur.

Del estudio que ha realizado George R. McGhee, Jr., de la Universidad de Rutgers, acerca de las faunas fósiles de Nueva York, que en el Devónico superior estaba cerca del ecuador, se deduce que la crisis duró en esa región unos siete millones de años. Este autor ha aportado también una observación interesante: mientras que otros grupos animales desaparecían, las esponjas silíceas se diversificaban; cuando la crisis remitió y los otros grupos se recobraban, el número de especies de esponjas

silíceas disminuyó. El sentido de esa observación es que, como ahora, la mayoría de las esponjas silíceas estaban adaptadas a vivir en aguas frías. Su prosperidad durante la crisis del Devónico superior constituye una prueba circunstancial de que sus competidores menos afortunados sucumbían a los efectos del enfriamiento climático.

Establecer una relación entre enfriamiento y crisis marinas del Paleozoico superior y del Mesozoico inferior resulta más problemático. Un largo episodio glacial del Pennsylvaniense (Carbonífero superior) y Pérmico inferior no parece que causara una extinción particularmente importante, pero hay pruebas claras de que la vegetación tropical siguió prosperando cerca del ecuador, señal de que el enfriamiento se dio sólo en latitudes elevadas. Las especies de esas latitudes probablemente hallaron refugio en los trópicos.



10. CORALES ARRECIFALES TROPICALES Y BRAQUIOPODOS experimentaron pérdidas importantes en la crisis del Devónico superior, hace unos 370 millones de años. Los corales devónicos (*izquierda*) nunca recobraron un papel significativo como constructores arrecifales. Los braquiópodos (*derecha*) abundaron a lo largo de todo el Paleozoico, aunque muchas familias de braquiópodos que desaparecieron en el Devónico superior quizás estuvieran limitadas a las aguas cálidas: el 90 por ciento de ellas no están representadas en la fauna fósil de las aguas frías de Sudamérica, situada entonces cerca del polo Sur.

Una extinción masiva marca la parte más moderna del período pérmico, pero siguió una pauta compleja. Los animales marinos polares adaptados al frío, que fueron los primeros en extinguirse, parece que sucumbieron a la expansión de los climas cálidos a partir de latitudes bajas: en el Hemisferio Sur, por lo menos, su extinción coincidió con la reducción de los glaciares polares y con una migración hacia el polo de plantas terrestres adaptadas al calor. Por el contrario, las extinciones que se dieron en los Trópicos más tarde, cerca del final del Pérmico, quizá fueran de la mano de un enfriamiento climático ampliamente extendido, a pesar de que las pruebas de tal enfriamiento sean hasta ahora solamente locales. En las sucesiones rocosas de la parte norte de la Unión Soviética, que son del Pérmico superior, se encuentran restos sedimentarios groseros, del tipo de los que pueden soltarse de bloques de hielo marino, en la parte alta de calizas, que tienden a ser depositadas en aguas más calidas. Esta secuencia sugiere una expansión de hielo marino en esa época. Además, en el Hemisferio Norte muchas de las especies marinas que sobrevivieron hasta el Triásico, primer período del Mesozoico, soportaban una amplia gama de temperaturas. Muchos de los ammonoideos del Mesozoico inicial, por ejemplo, vivían tanto en el Tethys como en aguas frías; e igualmente los conodontos, un grupo de animales nadadores con forma de anguila que atravesaron la crisis del Pérmico superior sin sufrir mayores pérdidas.

La información de que disponemos acerca de las extinciones masivas a fines del Triásico y del Jurásico, dispersa, no puede vertebrarse todavía. Parece que la crisis triásica eliminó los arrecifes coralinos tropicales de Europa meridional, mientras que el registro de las plantas fósiles sugiere que un enfriamiento notable en Europa y Asia septentrional acompañó las extinciones jurásicas. Sin embargo, faltan muchas pruebas para que pueda afirmarse que esos acontecimientos coincidieron con un enfriamiento climático de gran extensión.

Las especies animales más antiguas del océano actual son las que prosperan en aguas frías. Muchas especies árticas características, como algunos braquiópodos, estrellas de mar y bivalvos sedimentívoros, pertenecen a órdenes biológicos cuyos orígenes se remontan centenares de millones de años, hasta el Paleozoico. Por el contrario, faunas tropicales como las comunidades arrecifales, sometidas a extinciones masivas periódicas, han aparecido y desaparecido con gran rapidez a lo largo de los tiempos geológicos. La comunidad arrecifal, formada por corales y esponjas, diezmada en la crisis del Devónico superior, cedió el paso a una asociación menos eficaz de constructores de arrecifes que comprendía briozoos, braquiópodos, esponjas calcáreas y algas; después de que esta asociación sufriera una extinción generalizada a fines del Pérmico, fue sustituida, a su vez, por una comunidad de corales parecidos a los existentes actualmente. Los corales perdieron su hegemonía, al parecer a causa de los rudistas del Cretácico superior, y sólo llegaron a predominar cuando los rudistas fueron aniquilados por la crisis del Cretácico terminal. Adaptadas, a menudo, a una gama estrecha de temperaturas, y sin posible escape en tiempos de enfriamiento, las especies tropicales son las más dañadas por los cambios del clima.

Se ha afirmado con frecuencia que las especies tropicales estaban menos capacitadas para tolerar cambios en su ambiente. Si ello fuera cierto, las alteraciones fuertes de un factor limitante que no fuera la temperatura habrían producido extinciones masivas con sesgo tropical. Afortunadamente para la hipótesis del enfriamiento climático, no existe prueba alguna que indique que la mayoría de las especies tropicales marinas sean más especializadas en sus necesidades, excepto en relación con la temperatura, que las especies templadas. Las necesidades alimenticias valen de ejemplo. La mayoría de especies marinas no microscópicas son herbívoras; en cualquier latitud son más comedores universales que especializados, consumiendo muchos tipos de algas, bacterias o restos de plantas.

Al establecer la conexión entre cambio climático y crisis marina he tenido buenas razones para dejar de lado una pregunta obvia: ¿qué causó el enfriamiento climático que, a su vez, provocó las extinciones masivas? Se han sugerido varias explicaciones para el periódico avance de los glaciares polares registrado durante los últimos millones de años, entre las que se cuentan las fluctuaciones de la órbita terrestre. El enfriamiento ecuatorial que acompañó las crisis marinas más antiguas pudiera reflejar una reducción en la emisión de radiación solar –un cambio en la constante solar– pero también podrían hallarse implicados otros factores. Aunque un examen continuado del registro fósil aportara más pruebas de la relación entre enfriamiento y extinciones marinas, quizá nunca conozcamos las razones exactas por las que se produjeron esas variaciones climáticas hace millones o centenares de millones de años.

Ciencia y sociedad

Canarias, observatorio de Europa

La astrofísica española está atravesando por una etapa de asentamiento institucional que le permitirá establecer la infraestructura necesaria y las líneas de investigación acorde con la misma. La materia prima, unos picos de observación excepcionales en un firmamento limpio, nos ha venido regalada por la naturaleza.

Uno de los institutos astrofísicos mejor dotados de nuestro país es, sin duda, el de Canarias (IAC). Su difícil historia nos la resume el propio director Francisco Sánchez, conocido por los lectores de INVESTIGACIÓN Y CIENCIA ("Luz zodiacal", mayo, 1977). En 1856 el almirantazgo británico, a propuesta del Astrónomo Real, acordó enviar a Charles Piazzi Smythe a la isla de Tenerife, al objeto de estudiar el mejoramiento de la observación astronómica con la altura. Allí, "con la aprobación de las autoridades españolas", cuenta Smythe, "siempre dispuestas en aquella isla a favorecer los propósitos de los científicos de cualquier país", acampó en Guajara (2717 metros de altura) y en Altavista (3250 metros), pudiendo, en efecto, comprobar las ventajas de los observatorios de montaña. Informó de sus resultados a la Royal Society. En 1858 publicaba su libro *Tenerife: An Astronomer's Experiment*; Tenerife se convertía en el precursor de los modernos observatorios de montaña.

OBSERVATORIO DEL TEIDE

DIAMETRO (CM)	DESCRIPCION	PROPIETARIO	OPERATIVO (AÑO)
30	FOTOPOLARIMETRO DE FUENTES DIFUSAS	OBS. BURDEOS (F)	1964
25	HELIOGRAFO	IAC (E)	1969
40	TELESCOPIO SOLAR AL VACIO	INST. KIEPENHEUER (A)	1972
155	COLECTOR INFRARROJOS	IAC (E)	1972
50	PROPOSITO GENERAL	UNIV. MONS (B)	1972
6 × 20	FOTOPOLARIMETRO 6 CANALES	IAC (E)	1984
80	PROPOSITO GENERAL (CONS. IAC)	IAC (E)	1984
45	TELESCOPIO SOLAR AL VACIO	UNIV. GÖTTINGEN (A)	1985
60	TELESCOPIO SOLAR VERTICAL	INST. KIEPENHEUER (A)	1985
90	THEMIS	I.N.A.G. (F)	EN NEGOCIACIONES

(A) = República Federal de Alemania, (B) = Bélgica, (F) = Francia, (E) = España.

1. El observatorio del Teide se halla instalado a 2400 metros de altitud en la zona de Izaña, en la isla de Tenerife; mide una superficie de 50 hectáreas.

OBSERVATORIO DEL ROQUE DE LOS MUCHACHOS

DIAMETRO (CM)	DESCRIPCION	PROPIETARIO	OPERATIVO (AÑO)
60	PROPOSITO GENERAL	REAL ACAD. CIENC. (S)	1982
60/44	HELIOSTATO/CASSEGRAIN	REAL ACAD. CIENC. (S)	1982
100	ASTROMETRICO	SERC (GB)	1984
250	PROPOSITO GENERAL («ISAAC NEWTON»)	SERC (GB)	1984
420	PROPOSITO GENERAL («WILLIAM HERSCHEL»)	SERC (GB)	1986
–	CIRCULO MERIDIANO	UNIV. COPENHAGUEN (D)	1984
350	PROPOSITO GENERAL	ITALIA	EN NEGOCIACIONES
250	PROPOSITO GENERAL	NORDICO	EN NEGOCIACIONES

(D) = Dinamarca, (GB) = Gran Bretaña, (S) = Suecia.

2. El observatorio del Roque de los Muchachos se halla a 2400 metros de altitud en la zona más alta de la isla de La Palma; su superficie es de 189 hectáreas.

Varios astrónomos bajaron el siglo pasado a Tenerife: Simony y Angstrom, entre ellos; en nuestro siglo, C. Müller, Ekron y Jean Mascart. Este último, del observatorio de París, quedó tan satisfecho de las condiciones de la isla que propuso la creación de un observatorio internacional en la montaña Guajara, iniciándose con este motivo una larga correspondencia entre los gobiernos alemán y francés por un lado y español por otro. Pero la primera guerra mundial acabó con la idea.

Con motivo del eclipse de sol de 1959, visible como total en Canarias, son muchos los astrónomos de todo el mundo que visitaron las islas. Vuelve a surgir la idea de instalar un observatorio permanente. Lentamente, y venciendo toda clase de dificultades, nace y va creciendo y consolidándose un grupo de astrofísicos ligado a la Universidad de La Laguna. Los años sesenta se gastan en trabajos de prospección y de formación de astrónomos.

Por fin, el 10 de octubre de 1970 se inaugura el Observatorio del Teide. Tres años más tarde nace el Instituto Universitario de Astrofísica. En 1975 se constituye, por la acción coordinada de la Universidad de La Laguna, el Consejo Superior de Investigaciones Científicas y la Mancomunidad Interinsular de Cabildos de la Provincia de Santa Cruz de Tenerife, el primitivo Instituto de Astrofísica de Canarias. Prosigue la investigación nacional y menudean las visitas de observadores extranjeros. Se inicia el estudio de la zona del Roque de los Muchachos y Fuente Nueva, en la isla de la Palma. Canarias se convierte en la zona del planeta mejor estudiada en cuanto a su calidad astronómica. En 1979 se firma en Santa Cruz de la Palma el acuerdo de cooperación en materia de astrofísica entre España, Dinamarca, Gran Bretaña y Suecia. (En abril de 1983 se adhiere la República Federal de Alemania, y se internacionaliza también el Observatorio del Teide.)

Por ley 7/1982 del 30 de abril, se concedió al Instituto de Astrofísica de Canarias una especial personalidad jurídica al crear con esa denominación un consorcio público de gestión integrado por la administración del estado, el gobierno autónomo canario, la Universidad de La Laguna y el CSIC. Pero la

normativa legal tiene escaso valor si no facilita y fomenta, con la subvención necesaria, los programas que se den a sí mismos los propios centros. *Investigación y Ciencia* ha procurado acercar la labor de los astrónomos de Canarias, centrada especialmente en astronomía solar, a sus lectores ("Las oscilaciones solares", de Teodoro Roca Cortés, mayo, 1982; "Las fulguraciones solares", de J. I. García de la Rosa, F. Herrera y M. Vázquez, diciembre, 1983).

El centro tiene esperanzas de que en un futuro inmediato los mayores telescopios del mundo (18 metros de diámetro y LEST) se instalen en sus observatorios. La inversión superará pronto, en la primera fase, los 10.000 millones de pesetas, de los cuales España aportará sólo alrededor de la décima parte (en obras de infraestructura y servicios), recibiendo como contrapartida principal el 20 por ciento del tiempo de observación libre de gastos, más el 5 por ciento para programas cooperativos.

Se ha previsto inaugurar oficialmente el Instituto y su Observatorio del Roque de los Muchachos hacia mediados del año próximo. Podrá decirse entonces que Europa mira al cielo en Canarias: Alemania, Dinamarca, España, Gran Bretaña y Suecia, como firmantes del acuerdo; Holanda e Irlanda participan a través del Reino Unido; Noruega, Finlandia, Italia y Francia han iniciado ya las negociaciones.

Ante semejante oportunidad, los propios astrónomos españoles convocaron a sus colegas europeos para evaluar el trabajo realizado, las líneas de investigación abiertas y los posibles cambios para un futuro más provechoso. Fruto de ello fueron las *Recomendaciones sobre el futuro de la investigación en el IAC*, publicadas por el propio centro. Se empieza por reconocer aquí el valor de lo realizado: "debería continuarse y fomentarse la investigación actual en física solar, en el infrarrojo y en instrumentación", los tres ámbitos más trabajados por los astrónomos españoles. Ponen el dedo en una de las llagas más extendidas por todo el cuerpo de nuestra investigación –astrofísica o no–: "El espacio previsto para la biblioteca en el nuevo edificio del IAC debería aumentarse considerablemente, ya que una biblioteca adecuada es vital para un instituto de astrofísica importante".

Y un aviso para navegantes: la interacción entre la enseñanza universitaria de la astrofísica y la investigación. Nunca debería yugularse la docencia de los astrofísicos del IAC en sus centros o en La Laguna por mor de una normativa estrecha.

Alfonso X el Sabio y la música

El siglo XIII es decisivo para la historia de la música en Occidente. Como las restantes manifestaciones del arte y de la ciencia, la música va inseparablemente unida en España a la persona del rey Alfonso X el Sabio. Durante su reinado no se produce ningún fenómeno enteramente nuevo, pero se dan pasos que harán cambiar, a los ojos del historiador del siglo XX, la fisonomía musical de la Península.

Hacía siglos que se habían puesto las bases para una nueva comprensión y un desarrollo coherente de la técnica polifónica. Los ingleses aparecen ante nosotros, ya en el siglo IX, como auténticos pioneros de un arte nuevo ligado a la invención de los tropos. El canto llano que diariamente se practicaba en los oficios sagrados de las iglesias monásticas y catedrales era, en opinión de uno de los principales creadores de tropos –el monje Notkero (h.850)–, aburrido y monótono. Si al mismo se le intercalaba un texto o una melodía, se obtenía una nueva forma que hacía variar la monotonía del canto llano y abría el camino a múltiples posibilidades de creación musical, entre las que se contaba la polifonía. Mientras el coro de monjes o de clérigos (coro reducido, por cierto, a los más avezados en la técnica musical) interpretaba los prolijos melismas del canto llano, el maestro cantor se arrancaba desde esos mismos neumas con una nueva melodía produciendo un efecto inédito de consonancias que convertían la superficie plana del canto gregoriano –después que éste había perdido ya su primitiva riqueza y relieve– en la perspectiva arquitectónica del discanto polifónico.

En la segunda mitad del siglo XIII la técnica polifónica aparece ya bastante depurada. Los primeros *órganos* y conductos de la Escuela de San Marcial de Limoges, entre los que destaca la colección del Códice Calixtino de Compostela, van dejando progresivamente lugar a los *motetes*, que son algo así como tropos de tropos, esto es, glosas que se intercalan y superponen al texto litúrgico en la melodía del discanto polifónico.

París se había convertido en centro musical por excelencia de toda la cristiandad. Su universidad albergaba a nobles castellanos y de otros reinos. Un hermano del futuro rey Alfonso X, Don Felipe, sería enviado a esta ciudad a estudiar el *trivium* y el *quadrivium* para poder ejercer dignamente los altos cargos eclesiásticos para los que estaba destinado. De allí llegó a la Península la técnica polifónica más avanzada. Digno testimonio de ello es el códice de las Huelgas, además de los fragmentos conservados en la zona de Burgos y en otras partes de Castilla y Cataluña.

La Corte andariega del rey Alfonso X no pudo prescindir de la polifonía litúrgica como instrumento de boato y solemnidad en las fiestas más importantes del año, allí donde se encontraba. Al lado de las capillas polifónicas que podríamos considerar como más estables en las grandes ciudades, como Sevilla, Toledo y Burgos, el rey llevaría probablemente en su Corte cantores y ministriles para las celebraciones más solemnes durante el camino.

Siglo y medio antes de la subida al trono del rey Alfonso X, las cortes europeas de reyes y nobles habían visto nacer un movimiento centrado en una actividad poético-musical, cuyo tema principal era el *fin'amor* del poeta músico, o trovador, hacia una dama, generalmente de noble linaje. Esta actividad era ejercida así por los más altos como por los más humildes personajes de la corte. Algunos de los reyes y príncipes fueron egregios trovadores, tales como Guillermo IX, duque de Aquitania, considerado el iniciador del movimiento trovadoresco, Teobaldo I, rey de Navarra, Alfonso II y Pedro el Grande, ambos de Aragón.

Las cortes de los reyes de Castilla habían sido tradicionalmente muy generosas y hospitalarias con los trovadores del otro lado de los Pirineos. De éstos aprendió el propio rey sabio, cuando era todavía príncipe, el arte de trovar. Diversos textos del final de su vida nos ilustran sobre su condición de trovador enamorado que cantó a su Dama con las más elevadas *razones* del *fin'amor*. Sabemos incluso de sus amores extramatrimoniales, más allá de esa mera *razón* literaria, de los que nacieron algunos de sus hijos más amados, como la futura reina de Portugal, Doña Beatriz. Ya anciano, desengañado del amor humano, seguirá sintiéndose trovador, pero ahora ya *a lo divino*. Y con este sentimiento compuso, o mandó componer, las Cantigas de Santa María, en las cuales la Virgen es la verdadera Dama, «que eu tenno per Sennora», dirá en una de ellas.

La música cumple en estas cantigas una función primordial, pues a través de ella se expresa el *loor* a María, característica de todos estos poemas al-

fonsinos, aun los que narran milagros y hechos portentosos.

Para cantar las Cantigas, así profanas como religiosas, la Corte real disponía de juglares y ministriles, personal cualificado que realizaba con dignidad, escenificando y danzando a veces, estos poemas musicales. Sabemos que en ciertas ceremonias más solemnes –aquéllas, por ejemplo, en que se armaba caballero a un personaje noble, unas lindas doncellas salían al medio de la iglesia para danzar mientras cantaban unas cantigas.

Hemos hablado de los juglares, ministriles y danzantes como intérpretes de la música. Eran éstos personas dotadas de gran habilidad para el canto, el uso de instrumentos y la danza. El instrumentario musical de la época alfonsina era muy variado. De él nos ha quedado una extraordinaria muestra iconográfica –sin duda, la más importante de toda la Edad Media– en los códices que nos transmiten las Cantigas de Santa María.

No hace falta imaginarse a los ministriles, cantores y danzantes interpretando una cantiga o discantando un *órgano* polifónico. Las miniaturas de los referidos códices nos representan, con el realismo que permite la abstracción iconográfica del arte gótico, numerosas escenas con la actuación de estos personajes.

El carácter lúdico que inevitablemente asociamos a estas representaciones no debe hacernos olvidar, sin embargo, la gravedad de la disciplina musical en los estudios generales, institución predecesora de la universidad. Era una de las cuatro asignaturas del *quadrivium*, cuyo objetivo delimita el rey sabio en su *General e Grande Estoria*.

El severo tratado sobre la música de Boecio era leído por hombres de gran prestigio. Y autores nuevos, como el franciscano fray Gil de Zamora, escribieron nuevas obras para resumir y poner al alcance del estudiante la difícil doctrina de los antiguos, utilizando incluso –no es el caso del religioso– libros de grandes autores árabes, como el *Catálogo de las Ciencias* de Al-Farabi.

En el ambiente clerical y universitario aparecían, de vez en cuando, personajes extraños que se ganaban el sustento con bufonadas y representaciones histriónicas del peor gusto, detrás de un lenguaje, culto y ramplón a la vez, del bajo latín eclesiástico, que sólo podía ser comprendido por los estudiantes. Los goliardos, así se llamaban estos auténticos *ganapanes*, visitaban las universidades más importantes de Europa, siendo Oxford, París y Bolonia el eje más recorrido.

El reinado de Alfonso X el Sabio, desgraciado en lo político, económico y hasta en lo familiar y dinástico, se nos revela como un período extraordinario para las ciencias y las artes. El historiador de la música debe reconocer en él, quizás, uno de los momentos más gloriosos. Nunca como entonces aparecen las artes más ligadas entre sí. Los hombres del primer gótico hispánico supieron recoger una herencia múltiple y dispersa para presentarla a sus sucesores con la unidad y coherencia que hizo posible el arte nuevo (*ars nova*) de los grandes creadores, músicos y poetas, del *trecento*. (Ismael Fernández de la Cuesta.)

Agente doble

El reciente anuncio del aislamiento, por parte de investigadores del Instituto Nacional del Cáncer de los Estados Unidos (NCI), del virus responsable del SIDA (síndrome de inmunodeficiencia adquirida) ha ido acompañado de síntomas de que el logro trasciende el ámbito científico, para incidir en el político y social. Al comunicado oficial precedieron rumores contrapuestos; el Secretario de Sanidad y Servicios Humanos y sus ayudantes asistieron a la conferencia de prensa para pedir confianza en el personal científico de la administración; en la prensa se ha aireado una controversia internacional sobre precedencia científica.

La publicación, en *Science*, de los correspondientes trabajos demuestra que, en efecto, un grupo del NCI encabezado por Robert C. Gallo ha culminado una laboriosa serie de investigaciones con el aislamiento de un nuevo virus, denominado HTLV-III, y su correlación convincente con la patogénesis. El aparente aislamiento de ese mismo virus en el Instituto Pasteur de París, el año pasado, sin que se reunieran pruebas de que provocaba el SIDA, no resta mérito al logro del NCI, ni, por supuesto, al francés.

La aparición súbita de una enfermedad devastadora, no conocida antes y de causas ignoradas, es un fenómeno raro en la sociedad avanzada del siglo XX. El SIDA lo ha sido. Apareció en los Estados Unidos a principios de 1981, y desde entonces se ha contabilizado 5000 casos, de los cuales el 45 por ciento con desenlace fatal. La enfermedad se presentó como un síndrome, grupo de síntomas, cuyo origen común era un defecto grave de la respuesta inmune. Los pacientes mostraban una carencia importante de linfocitos *T*, en particular de células *T* facilitadoras (*helper*), que, en presencia de antígenos extraños, instan la secreción de anticuerpos por parte de los linfocitos *B*. Los afectados sufrían numerosas infecciones, a menudo de patógenos poco virulentos, que no resultan peligrosos en personas con un sistema inmunitario entero.

En un principio pareció que el SIDA atacaba sólo a homosexuales de prácticas promiscuas o a varones bisexuales. Siguen constituyendo, con mucho, el grupo más afectado, pero corren riesgo también sus compañeras sexuales femeninas; los haitianos instalados en los Estados Unidos y los emigrantes de Africa ecuatorial que viven en Europa; quienes se administran drogas por vía intravenosa; los hemofílicos tratados con preparados obtenidos de sangre de bancos sanguíneos; los receptores reiterados de transfusiones de sangre y los hijos de padres incluidos en grupos de elevado riesgo. Al propio SIDA precede a menudo un pre-SIDA, un "pródromo" caracterizado por fatiga y trastornos de los ganglios linfáticos.

El virus responsable se había perfilado ya en una investigación emprendida hace 14 años. Se acababa de demostrar entonces que los virus causantes de ciertas leucemias animales y otros cánceres eran retrovirus, es decir, virus que emplean ARN como material genético, que se "retro-transcribe" a ADN en la célula animal infectada por medio de una enzima vírica, la transcriptasa inversa. Gallo y sus colaboradores emprendieron, en el NCI, el desarrollo de técnicas de detección de retrovirus que pudieran estar asociados a leucemias y linfomas humanos. Gracias a su descubrimiento del factor de crecimiento de las células *T*, proteína que insta el crecimiento a largo plazo de los linfocitos *T* en un medio de cultivo, y el desarrollo de un método fino de estimación de la transcriptasa inversa, se ha logrado identificar una familia de virus de leucemia y linfoma de células *T* humanas (HTLV), que constituyen los primeros retrovirus humanos y los primeros virus aparentemente responsables de cánceres humanos. HTLV-I se ha demostrado asociado a un cáncer de las células *T* endémico en la porción meridional de Japón, el Caribe y Africa. Un virus emparentado con él se denominó HTLV-II.

Gallo y sus colaboradores intuyeron que el agente causal del SIDA podía ser un miembro de la familia HTLV. Los in-

dicios eran numerosos. Los virus tendían a infectar células *T*, en especial las células facilitadoras. Otro retrovirus, el de la leucemia de los felinos, suprime la respuesta inmune, además de provocar leucemia con origen en las células *T*. Se conocía la incidencia de HTLV-I entre los caribeños (especialmente los haitianos) y africanos, y se había advertido que la transmisión estaba relacionada con los contactos íntimos y con los productos sanguíneos.

El año pasado, el grupo de Gallo aportó pruebas de la presencia de un HTLV en unos pocos pacientes del SIDA; lo mismo advirtieron Myron Essex y sus colegas, de la Escuela de Salud Pública de la Universidad de Harvard. La prueba no era, sin embargo, suficientemente clara para establecer en él la causa de la enfermedad. Según parecer de Gallo y sus colaboradores, debía tratarse de alguna variante de HTLV, distinta de HTLV-I y de HTLV-II, cuyo efecto principal no fuera tanto conferir inmortalidad a las células *T* cuanto lesionarlas. Más de una vez dieron con una variante así, pero su aislamiento era fugaz, pues las células afectadas morían antes de que de ellas pudiera aislarse el virus citopatógeno. Al cabo, Mikulas Popovic y sus colaboradores, del grupo de Gallo, encontraron una línea de células *T* cancerosas susceptibles de infección con la nueva variante citopatógena y que sobrevivían y proliferaban; se obtuvo así una población inmortal de células en la que podía propagarse el virus en abundancia.

Se comprobó que el nuevo virus era un miembro indudablemente distinto de la familia HTLV, por lo que se le denominó HTLV-III. Al igual que los HTLV-I y II, infecta las células *T* facilitadoras, pero no insta su proliferación, sino que las mata. En las micrografías muestra gran semejanza con los restantes miembros de la familia, pero su denso núcleo tiene aspecto inconfundiblemente cilíndrico.

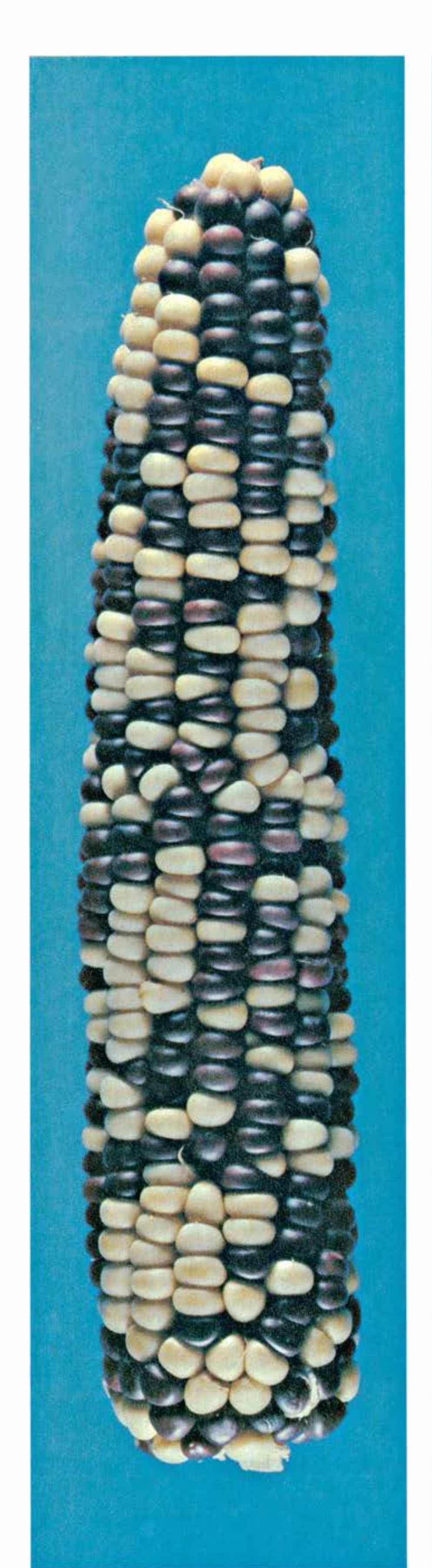
El paso siguiente que dieron los investigadores fue determinar presencia de HTLV-III entre los pacientes del SIDA. Se halló el virus en 18 de 21 muestras sanguíneas de pacientes de pre-SIDA, y en 26 de 72 muestras de pacientes del SIDA. Quizá se subestimara la presencia del virus, porque muchas de las muestras se recibieron en malas condiciones y porque la población de células *T* se había agotado en los pacientes que habían contraído la enfermedad tiempo atrás. Sólo se detectó el agente en una muestra de 22 varones sanos, homosexuales no promiscuos; seis meses después, el hombre que dio la muestra positiva manifestó la enfermedad.

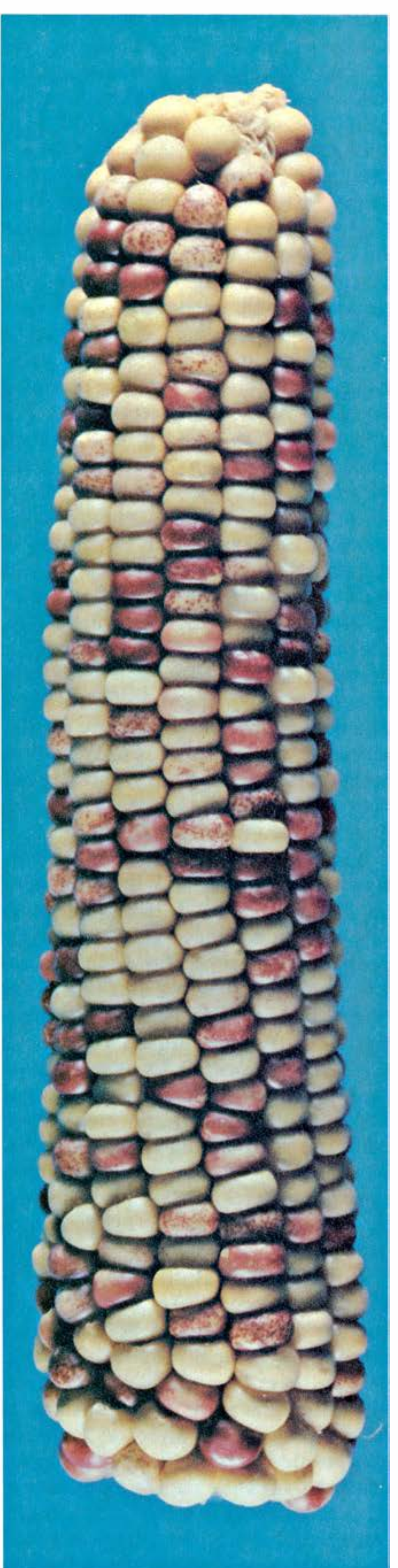
Aún sin llegar al aislamiento del virus, puede determinarse si una persona ha quedado expuesta a él por la presencia, en su sangre, de anticuerpos estimulados por los antígenos víricos. El grupo de Gallo encontró anticuerpos contra HTLV-III en el suero sanguíneo del 87,8 por ciento de un grupo de pacientes del SIDA y en el 78,6 por ciento de pacientes de pre-SIDA.

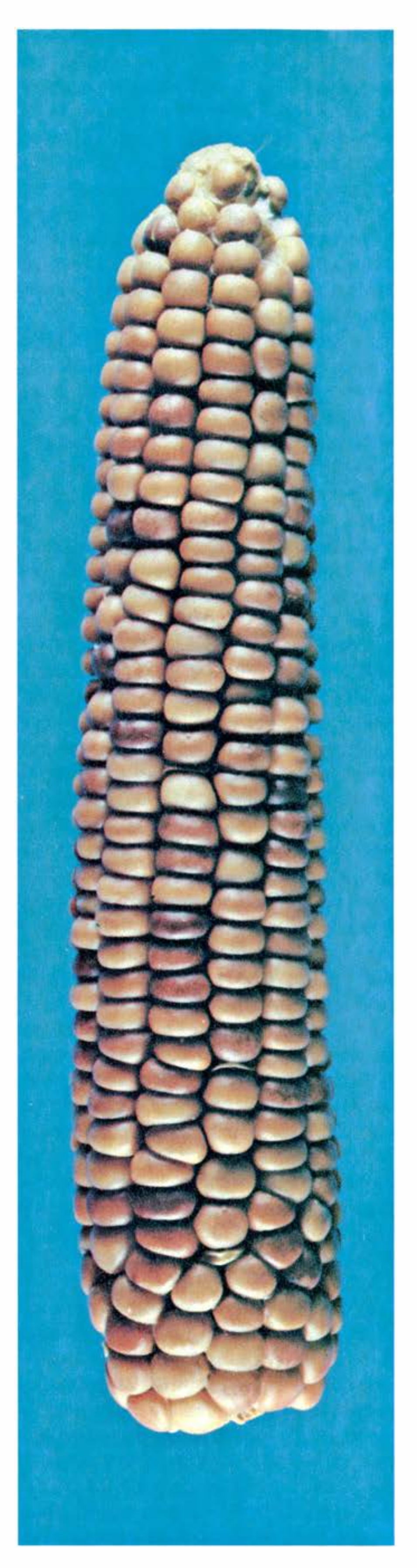
A partir de sus datos, Gallo y colaboradores concluyen que "HTLV-III es la causa primaria del SIDA". Los datos franceses también parecen apuntar a ello, puesto que es muy probable que el virus que aislaron y HTLV-III sean el mismo, o cuando menos miembros de subgrupos estrechamente emparentados. Parece que se ha exagerado la cuestión de la precedencia. Los autores franceses, encabezados por Luc Montagnier y Jean-Claude Chermann, lograron antes la identificación del virus, pero no pueden atribuirse el establecimiento de su papel causal; los estadounidenses publicaron su informe un año después, pero habían propagado el virus, lo habían descrito en detalle y encontraron el virus y anticuerpos contra él en muchos más pacientes.

En 1983, el grupo del Instituto Pasteur informó en *Science* del aislamiento, en un paciente de pre-SIDA, de lo que parecía un nuevo virus, no emparentado con HTLV-I. Lo denominaron LAV (por *lymphadenopathy-associated virus*, virus asociado a una linfadenopatía). A partir de entonces hallaron el LAV en la mayoría de los pacientes del SIDA explorados; otros investigadores franceses han encontrado anticuerpos contra el virus en una elevada proporción de los efectos de pre-SIDA estudiados, en un tercio de los pacientes del SIDA y sólo en uno de 130 donantes sanos de sangre.

El descubrimiento de un agente infeccioso sólo constituye un paso en el camino hacia el control de una enfermedad. El principal beneficio que cabe esperar de la identificación del HTLV y de su multiplicación en abundancia es la posibilidad de someter la sangre a ensayos de la presencia de anticuerpos contra el virus. Se cuenta también con desarrollar una vacuna profiláctica que estimule la síntesis de anticuerpos contra el virus. Sin embargo, no se tiene la certeza de que pueda estimularse una respuesta inmunológica suficientemente potente para que desarme un virus cuyo objeto principal es, precisamente, el propio sistema inmune.







1. MUTACIONES ESTABLES E INESTABLES afectan a la pigmentación de las semillas que estudió la autora. Una mutación estable en un gen necesario para la síntesis de pigmento púrpura explica la pérdida de color de algunas semillas de la mazorca de la izquierda. Las otras dos mazorcas son ejemplos de estructuras de variegación producidas por mutaciones inestables causadas por elementos genéticos transponibles. Cuando en los inicios del desarrollo de las semillas, o ya avanzado éste, dos versiones diferentes del elemento *Spm* abandonan por transposición un locus genético que afecta la síntesis de un pigmento púrpura, se producen manchas pigmentadas irregulares o pequeñas en semillas incoloras (*centro*). Las semillas "bronce" (*derecha*) débense a una mutación, producida por el elemento *Ac*, que interfiere la síntesis de pigmento. Cuando *Ac* abandona el gen, se producen sectores púrpura.

Elementos genéticos transponibles del maíz

Se han identificado en bacterias, plantas y animales genes móviles que Barbara McClintock descubriera hace treinta años. Varios de ellos se han caracterizado en el maíz hasta el nivel molecular

Nina V. Fedoroff

Un elemento genético transponible es un trozo de ADN que puede desplazarse de un sitio a otro del genoma (dotación total de material genético) de un organismo. Se escinde de una posición y se inserta en otra, sea en el mismo o en diferente cromosoma. Puede tener estructura simple y estar formado por sólo unos pocos genes que promueven la transposición, flanqueados por secuencias especiales de reconocimiento. El movimiento de un elemento transponible puede dar origen a mutaciones o reestructuraciones cromosómicas y afectar con ello a la expresión de otros genes. Se han descrito cromosomas salpicados de estos elementos móviles, o "genes saltarines", y se sospecha que su capacidad para modificar la expresión y estructura de otros genes, quizá hasta la estructura de un genoma entero, constituye un mecanismo importante que interviene en el cambio genético a largo plazo, punto crucial de la evolución orgánica.

Descubrimiento y aislamiento

Los elementos transponibles se descubrieron en plantas de maíz en la década de 1940, fruto de los notables estudios genéticos realizados por Barbara McClintock en el departamento de genética de Cold Spring Harbor, Nueva York, de la Institución Carnegie de Washington. En los últimos 20 años se han identificado elementos similares en todos los tipos de organismos examinados, desde bacterias hasta gusanos y moscas del vinagre. Algunos elementos se han aislado y analizado en detalle molecular, pero hasta hace muy poco los elementos del maíz han constituido una mera abstracción genética.

Durante los últimos años, en mi laboratorio del departamento de embriología de la Institución Carnegie de Baltimore, he tenido el placer de aplicar las nuevas técnicas de la biología molecular al aislamiento y estudio de los genes móviles del maíz definidos por McClintock. Algunos, descubiertos hace casi cuatro décadas, se han aislado ahora, y los análisis de su estructura y secuencias de ADN han permitido profundizar en la comprensión de cómo operan esos elementos.

Un nuevo enfoque

La conciencia de la abundancia y omnipresencia de los elementos transponibles es fenómeno reciente. Hasta finales de los años 70, el genoma de un organismo se consideraba estable. Los genes que determinan caracteres estructurales y bioquímicos son elementos discretos dispuestos a lo largo de cromosomas lineales. Cada gen ocupa una posición fija, que generalmente no cambia de una generación a la siguiente. La estabilidad de los cromosomas y de los genes presentes en ellos es la base del trabajo de los genéticos clásicos, que identificaron genes y determinaron sus posiciones cromosómicas sin aislar nunca un gen ni conocer su naturaleza química. Definieron un gen siguiendo la herencia de un carácter gobernado por un solo "factor" mendeliano hereditario. Cartografiaron las posiciones relativas de los genes mediante análisis de las descendencias de múltiples cruzamientos, que revelaron que los genes estaban ligados unos con otros en los cromosomas, y mostraron lo estrecho de su unión. Más tarde, genéticos y biólogos moleculares descubrieron el proceso por el que el material genético, identificado como ADN, se reproducía fielmente de generación en generación. Seguía destacándose la constancia de los genes y de los cromosomas y las reglas metódicas que regían la herencia.

Pero no hay regla sin excepción. Aunque la mayoría de los genes permanecen en su sitio, algunos se mueven. Quienes se esfuerzan en comprender las reglas a veces prestan poca atención a las excepciones; y así los primeros indicios de que algunos genes (o loci genéticos, para usar un término más general) no permanecían siempre en el mismo lugar pasaron desatendidos. Las primeras sospechas de transposición, que en su momento no se supo comprender, derivaron del análisis de las mutaciones inestables. Una mutación es un cambio del ADN; suele comportar la inactivación de un gen. En la mayoría de los casos el daño es reversible: se repara espontáneamente por retromutación, o reversión, en quizás uno de cada millón de descendientes de un organismo mutante. Sin embargo,



2. ESTRIAS ROJO-ANARANJADAS en una semilla de maíz, producidas por una mutación inestable. El elemento transponible *Mp* impide la síntesis de pigmento cuando se inserta en un locus llamado *P*. En algunas células se desplaza del locus; la mutación revierte y da lugar a sectores pigmentados. La transposición de *Mp* ha sido estudiada por Irwin M. Greenblatt, de la Universidad de Connecticut, quien proporcionó la semilla.



3. SEMILLAS DE MAIZ que muestran el efecto de mutaciones y transposiciones. El locus *C* produce un factor necesario para la síntesis de un pigmento púrpura (*a*). La inserción del elemento *Ds* en el locus inactiva el gen, privando de color a las semillas (*b*). En presencia del elemento *Ac*, *Ds* abandona el locus en algunas células; se producen entonces sectores de células pigmentadas, que se traducen en semillas con manchas (*c*). El locus *waxi* codifica una enzima necesaria para la síntesis de la amilosa, que hace traslúcido el endospermo (*d*). El endospermo es opaco cuando *Ac* está inserto en el locus (*e*). La transposición de *Ac* produce variegación en la transparencia de los sectores: unos son traslúcidos y otros opacos (*f*). Puede insertarse un elemento *Spm* "deficiente", o incompleto, en un locus de enzima para un pigmento, de suerte que el gen quede parcialmente inactivo y la semilla presente una pigmentación pálida (*g*). Un *Spm* completo ejerce dos efectos: su función supresora inactiva el gen, privando de color a la semilla, y su función mutadora transpone el *Spm* incompleto, produciendo clones de células revertidas, es decir, totalmente pigmentadas (*h*). Si la transposición es infrecuente, las manchas escasean (*i*); si la transposición tiene lugar en los comienzos del desarrollo de las semillas, los clones revertidos son grandes (*j*). Una versión cíclica de *Spm* produce manchas pigmentadas en cuyo seno aparecen pequeños sectores incoloros (*k*). La rotura cromosómica, de frecuente ocurrencia en los puntos de inserción del *Ds*, puede producir estructuras con sectores de pigmentación pálida separados por estrechas líneas limitantes profundamente pigmentadas (*l*).

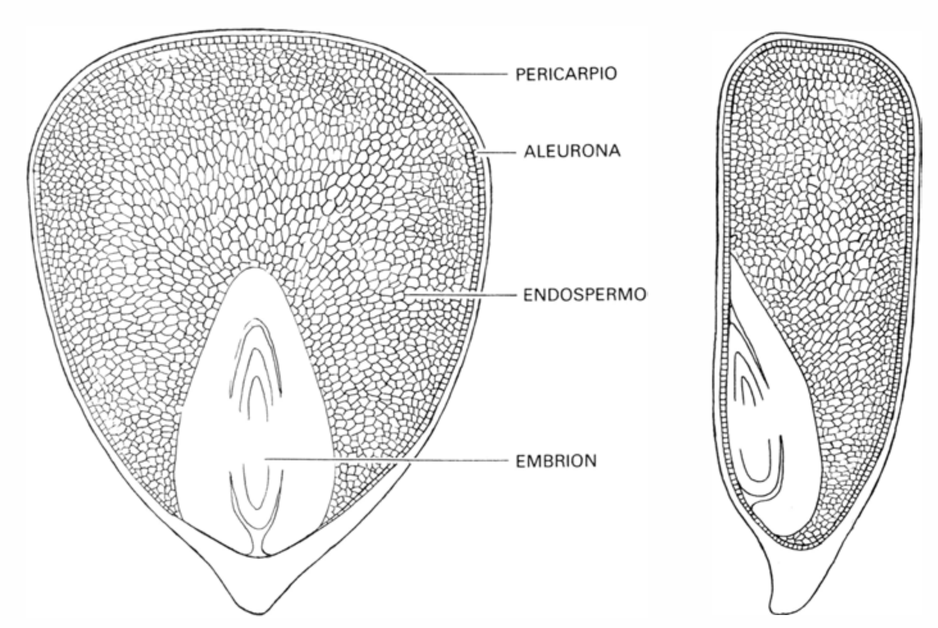
en unas pocas mutaciones, la frecuencia de reversión es tan alta que el organismo semeja un mosaico de tejidos mutantes y revertidos. Esas mutaciones inestables pueden conferir al organismo un aspecto variegado.

Variegación y reversión

En las primeras décadas de este siglo, R. A. Emerson, de la Universidad de Cornell, estudió un tipo de variegación muy familiar en las estirpes de maíz que se venden como maíz indio. La causa una mutación que interfiere la síntesis de un pigmento rojo anaranjado del pericarpio, la capa protectora de células que rodea la semilla de maíz. A diferencia de la mayoría de las mutaciones, ésta es tan inestable que revierte muchas veces durante el curso del desarrollo de cada semilla, dando lugar a estrías alternas de células pigmentadas y no pigmentadas. Emerson comprendió que las estrías se producían por mutaciones inestables, pero no supo apreciar lo que las causaba.

En los años 30, Marcus M. Rhoades, actualmente en la Universidad de Indiana, descubrió que la inestabilidad genética del maíz podía ser condicional. Lo que semeja una mutación estable puede repentinamente hacerse inestable en presencia de un gen determinado. La mutación que estudió desbarata la síntesis de un pigmento púrpura de la aleurona, la capa más externa del endospermo de la semilla (inmediatamente por debajo del pericarpio), haciendo incolora la aleurona. Rhoades demostró que si en la célula se encontraba un determinado gen, la mutación revertía, produciendo manchas de pigmento purpúreo en una aleurona incolora. Denominó *Dotted* (moteado) al locus que daba origen a esas manchas. No sabía, sin embargo, que el locus *Dotted* podía desplazarse.

Fue Barbara McClintock la primera en comprender que un elemento genético podía saltar (transponerse). El primer trabajo en donde planteaba que un locus genético podía desplazarse, publicado en 1947, produjo cierta sensación de excentricidad, y durante años nadie supo qué hacer con él. Se le reconoce ahora como uno de los descubrimientos genéticos más importantes del siglo; el año pasado se concedió a McClintock el premio Nobel por sus trabajos. Descubrió la transposición estudiando las propiedades de los cromosomas fragmentados del maíz. Tales cromosomas, según vio, originaban una maraña genética. Eran muy inestables y participaban en reorganizaciones cromosómicas que alteraban profundamente el genoma. Observó, además, que las plantas con cromosomas rotos a menudo producían descendencia variegada. La variegación afectaba a diversos caracteres, lo que sugirió a McClintock que se generaban numerosas mutaciones inestables y nuevas.



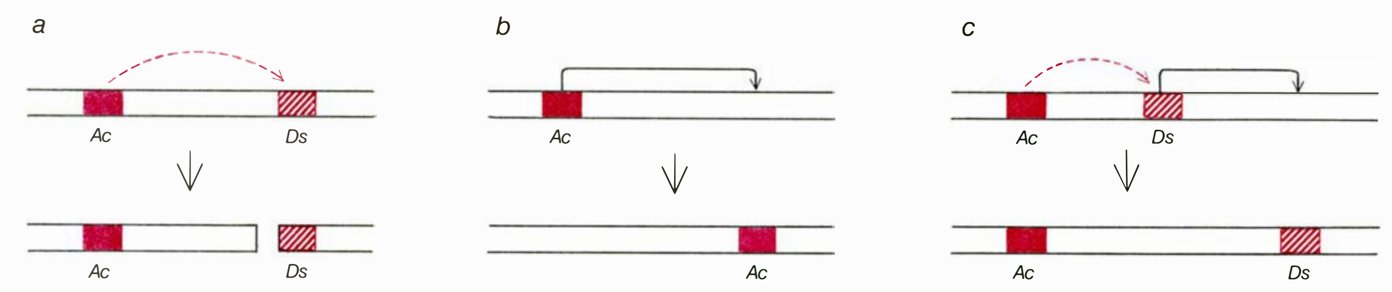
4. SEMILLA DE MAIZ en dos secciones. La componen el embrión de la futura planta y su tejido nutricio, el endospermo. Muchas de las mutaciones inestables analizadas con mayor detalle afectan a la síntesis de los pigmentos de los tejidos de aleurona y pericarpio (la capa protectora) o al almidón del endospermo.

El primer locus cuya transposición observó McClintock resultó ser en realidad un raro tipo de locus. Se manifestaba como un punto específico de rotura, o disociación cromosómica; lo denominó locus *Dissociation* o *Ds*. Aunque *Ds* constituye el punto de rotura, no es en sí mismo responsable de la rotura, que sólo se desarrolla si está presente otro locus; a ese segundo locus lo llamó *Activator*, o *Ac*, por su capacidad para instar la rotura en el locus *Ds*. Observó que el tipo de herencia de *Ac* y *Ds* no era el corriente. De ordinario, *Ac* actuaba como un locus convencional, pero en algunos casos (sólo en un pequeño porcentaje de la descendencia) desaparecía o se trasladaba a una nueva posición del mismo o de otro cromosoma. El locus *Ds* podía también transponerse, pero era incapaz por sí mismo de romper el cromosoma donde residía: no podía transponerse en ausencia de *Ac*. McClintock comprendió que *Ac* podía desplazarse por sí mismo (de forma autónoma), pero que *Ds* sólo se movía si lo activaba *Ac*.

Reversión y transposición

McClintock llevó a cabo una serie de experimentos que aclararon la relación que guardan las mutaciones inestables con la transposición. En uno de los casos que analizó, *Ds* había pasado de su posición original al locus llamado *C*, estudiado antes por varios genetistas. El locus *C* elabora un factor necesario para la síntesis de un pigmento púrpura de la aleurona; se sabía que las mutaciones de *C* interferían la producción de pigmento. McClintock recogió meticulosamente pruebas de que la nueva mutación de *C* se debía a la inserción de *Ds* en él. Comprobó que, en presencia de *Ac*, se producía en *C* el mismo tipo de rotura cromosómica que se había detectado en la posición original de *Ds*. Sin embargo, *Ds* tenía entonces un segundo efecto. Producía en *C* una mutación parecida a las estudiadas por Emerson y Rhoades. En ausencia de *Ac*, la mutación era estable: las semillas eran incoloras y toda la planta era verde. En presencia de *Ac*, en cambio, la mutación revertía en algunas células, con lo que tanto las semillas como la planta mostraban sectores de pigmento púrpura [*véanse las semillas* a-c *de la figura 3*].

McClintock se encontraba ya en condiciones de comprender la base genética de las mutaciones inestables. En una pequeña fracción de las semillas, la reversión había sido suficientemente precoz para afectar a las células geminales; podía, por tanto, transmitirse a las plantas de la descendencia. McClintock cultivó plantas a partir de las pocas semillas que habían revertido completamente y descubrió que, en todos los casos, *Ds* no se encontraba en el locus *C*. El locus funcionaba entonces normalmente. Además, incluso en presencia de *Ac*, el cromosoma no se rompía por el locus *C*, como ocurría cuando *Ds* se albergaba allí.



5. LOCUS *ACTIVADOR*, o *Ac*, como lo denominó Barbara McClintock por su capacidad para activar la rotura cromosómica de otro locus: el *Dissociation* o *Ds* (*a*). En presencia de *Ac*, la mutación provocada por *Ds* revertía. Esta figura muestra ambos loci sobre el mismo cromosoma, pero pueden hallarse en cromosomas diferentes. *Ac* es capaz de promover su propia transposición (*b*) o la de *Ds* (*c*) a otro punto del mismo o de otro cromosoma. *Ds* no puede desplazarse a menos que esté *Ac* en la misma célula. *Ac* es un elemento transponible autónomo y *Ds* es un elemento no autónomo de la misma familia.

McClintock concluyó que la inserción de *Ds* había causado la mutación original a aleurona incolora, y que la supresión de *Ds* era la responsable de la reversión. De estas observaciones dedujo que la variegación en plantas y semillas resultaba de la escisión de *Ds* del locus *C* en gran número de células durante el desarrollo de la planta. Su deducción se confirmó por la observación de que los sectores púrpura de las semillas no presentaban las pequeñas regiones incoloras que cabría esperar de haber persistido la rotura cromosómica asociada con el locus *Ds*. McClintock llegó a la conclusión de que una mutación inestable podía derivar de la inserción de un elemento transponible en un locus. El frecuente abandono del locus por parte del elemento transponible durante el desarrollo explica la reiterada reversión que caracteriza a las estructuras variegadas.

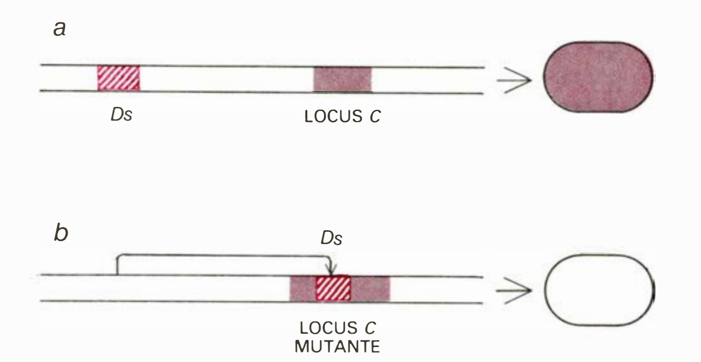
En los años que siguieron a su primera descripción de la transposición continuó McClintock su análisis de *Ac* y *Ds*, juntamente con otros pocos genetistas de maíz, identificando y caracterizando nuevos elementos. R. Alexander Brink, de la Universidad de Wisconsin en Madison, encontró que la mutación inestable estudiada por Emerson, que afectaba al pigmento rojo anaranjado del pericarpio, resultaba de la presencia de un elemento transponible del locus *P*. El elemento, al que llamó *Mp*, se demostró genéticamente idéntico a *Ac*. Del mismo modo que *Ac*, producía una mutación inestable insertándose en el locus *P* y abandonando el locus con elevada frecuencia.

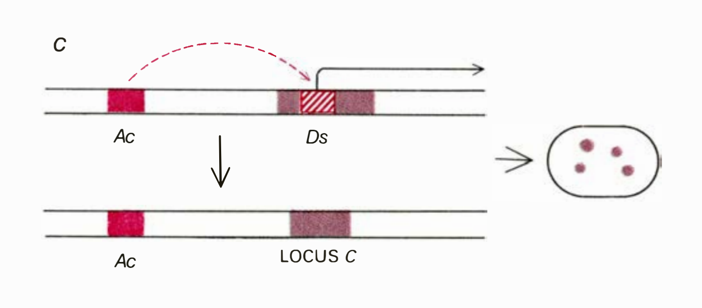
Brink y sus alumnos, especialmente Irwin M. Greenblatt, de la Universidad de Connecticut, se valieron de elegantes experimentos para aclarar cómo se transpone *Mp* (o *Ac*). Se desplaza durante la fase del ciclo celular en la que los cromosomas se replican para preparar la división celular. La transposición ocurre después de que el propio *Mp* se haya replicado, pero frecuentemente antes de que el cromosoma en el que reside haya completado su replicación. Solamente una de las dos copias hijas del elemento se desplaza; en la mayoría de los casos a una posición próxima del mismo cromosoma. Si el sitio receptor no se ha replicado todavía, su inmediata replicación da origen a un cromosoma hijo que transporta dos copias de *Mp* y un cromosoma con una sola copia en una nueva posición. Si el elemento se traslada a un sitio que ya se ha replicado, un cromosoma hijo no tiene *Mp* y el otro lleva dos copias del elemento, una de las cuales ocupa una posición nueva. La separación del elemento del locus *P* conlleva la reversión de la mutación inestable. El locus funciona a partir de entonces de forma normal, permitiendo la síntesis de pigmento rojo anaranjado.

Seguía sin aclararse la relación entre el locus *Ac* (o *Mp*) y el locus *Ds*. De los análisis genéticos de McClintock resultaba claro que los dos loci eran móviles, pero que la movilidad de *Ds* dependía de la presencia de *Ac* en el mismo genoma. McClintock observó que, en algunos casos, una mutación con las propiedades de *Ac* cambiaba a otra con el carácter de mutación *Ds*. Sospechó que los loci *Ac* y *Ds* guardarían alguna relación. La naturaleza física de los dos loci y su relación precisa no llegaron a aclararse, sin embargo, hasta el año pasado, cuando se aislaron.

Transposones en bacterias

En las décadas que separan la identificación genética de *Ac* y *Ds* y su aislamiento molecular se descubrieron elementos transponibles en otros organismos; comenzaba a desentrañarse la base molecular de su transponibilidad. La primera sospecha de que había elementos transponibles en las bacterias se tuvo a mitad de los años 60; se descubrió entonces que ciertas mutaciones peculiares no respondían a pequeños cambios en la secuencia de nucleótidos (las cuatro subunidades de ADN cuya secuencia codifica la información genética), sino a la inserción, en un gen, de una pieza de ADN extraño de tamaño apreciable. Se dedujo entonces que la resistencia bacteriana múltiple y trans-





6. SE PRESENTA UNA MUTACION EN EL LOCUS *C*, gen necesario para la síntesis de un pigmento púrpura en la aleurona (*a*), cuando *Ds* se inserta en el locus (*b*). La mutación inutiliza el gen; el pigmento no se sintetiza y la aleurona aparece incolora. Sin embargo, ante la presencia de *Ac* en el genoma, *Ds* abandona por transposición el locus en algunas células durante el desarrollo de las semillas (*c*). La mutación revierte cuando el elemento se desplaza, produciendo células en las que el locus *C* vuelve a ser funcional. Cada célula de ésas da origen, a su vez, a un sector de aleurona pigmentado.

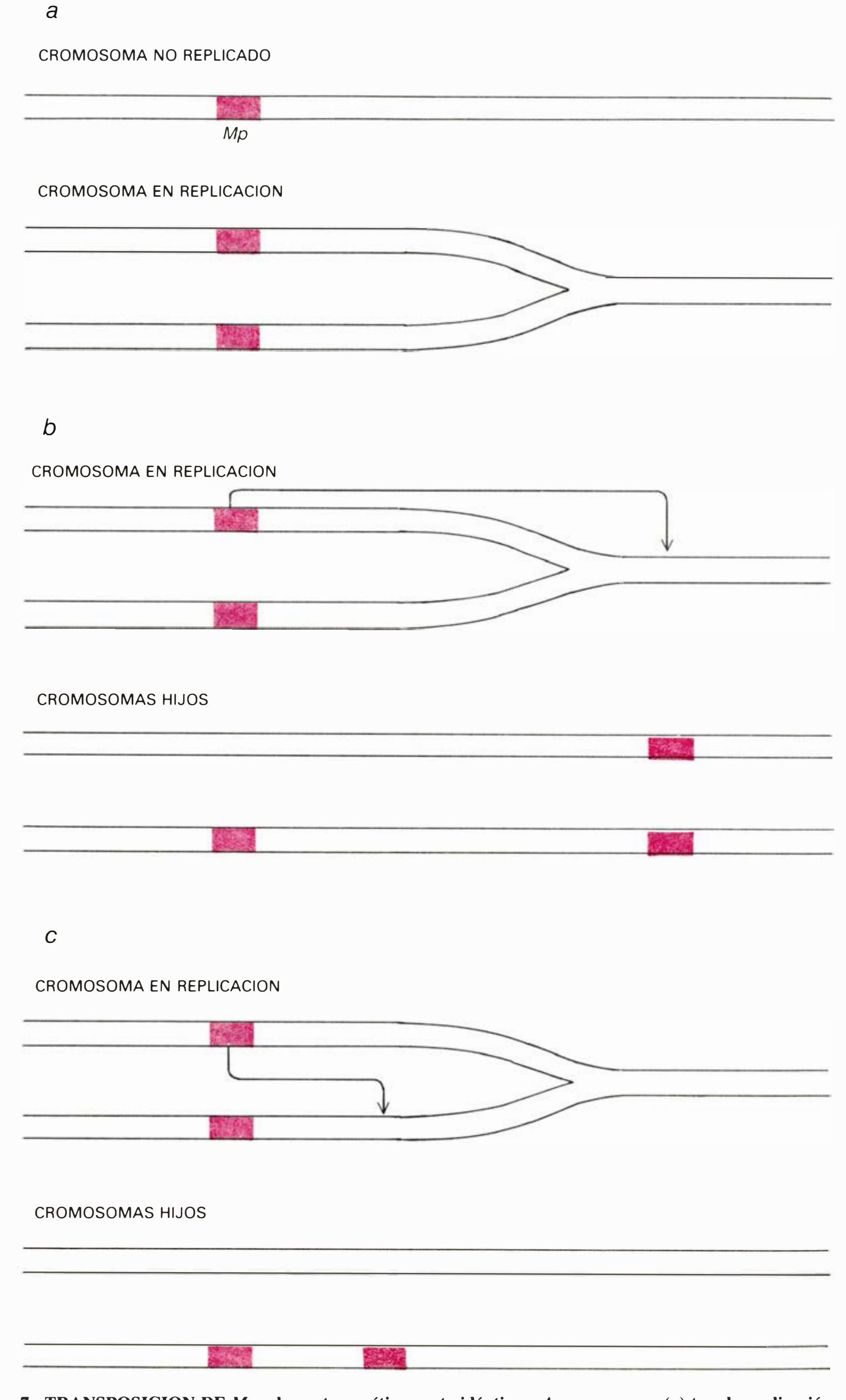
ferible a los fármacos se debía a elementos genéticos móviles. El amplio uso de antibióticos en animales y seres humanos había fomentado la aparición de estirpes de bacterias resistentes, simultáneamente, a varios antibióticos. Los microbiólogos descubrieron que los genes que conferían resistencia se transferían de una estirpe a otra por medio de pequeñas moléculas anulares de ADN bacteriano llamadas plásmidos. Hacia mediados de los años 70 se dedujo que los genes que conferían resistencia a antibióticos eran, en realidad, viajeros en pequeños elementos móviles, llamados transposones, que los plásmidos acarreaban en su tránsito de una célula a otra.

Algunos investigadores comprendieron inmediatamente que los transposones bacterianos y los elementos transponibles de las plantas de maíz se asemejaban más de lo que diferían [véase "Elementos genéticos transponibles", por Stanley N. Cohen y James A. Shapiro; INVESTIGACIÓN Y CIENCIA, abril de 1980]. Mas, para la mayoría, el trabajo de los genetistas del maíz sobre loci transponibles parecía confuso y de dudosa importancia para la moderna biología molecular. Los elementos bacterianos sí se analizaron en detalle y llegaron a comprenderse algunos de los mecanismos de su movilidad.

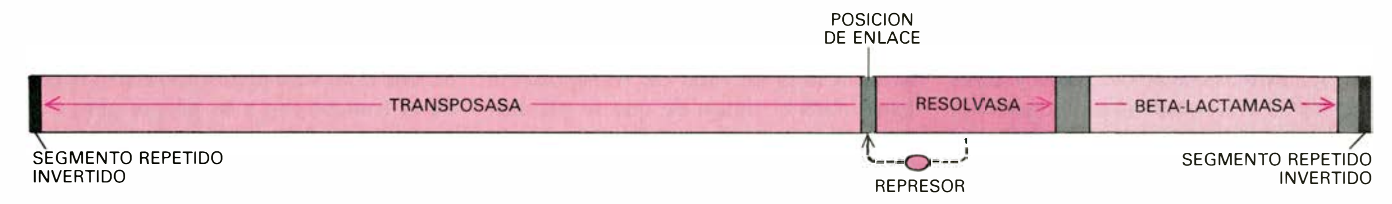
Quizás el transposón mejor estudiado sea el denominado Tn3. Mide unos 5000 nucleótidos de longitud y lleva tres genes. De éstos, dos codifican enzimas necesarias para la transposición; el tercero es el gen viajero que determina la enzima beta-lactamasa, que inactiva la ampicilina y confiere resistencia al antibiótico a las células bacterianas que portan el Tn3. Los extremos del transposón Tn3 son secuencias especiales no codificadoras, llamadas repeticiones invertidas. (Los cuatro nucleótidos, *A*, *G*, *T* y *C*, se enlazan en pares complementarios para formar la doble hélice de ADN: *A* se aparea con *T*; *G* con *C*. Las secuencias de los extremos de un transposón son simétricas bidireccional y rotacionalmente, esto es, la misma secuencia aparece en ambos extremos del elemento, pero se lee en dirección opuesta en los filamentos opuestos.) Las repeticiones invertidas sirven de señales de reconocimiento que delimitan la secuencia que ha de transponerse. En la transposición intervienen dos enzimas, una "transposasa" y una "resolvasa". La transposasa comienza el proceso de transposición; la resolvasa lo termina.

Al igual que los elementos del maíz, los transposones bacterianos producen mutaciones por inserción. Toda mutación que afecte a un gen esencial será letal para la célula, por lo que una frecuencia muy elevada de transposición aumentará el riesgo de que se produzcan tales mutaciones. El transposón dispone, por ello, de un mecanismo para regular la expresión, tanto de los genes de la transposasa como de la resolvasa. La propia resolvasa es el regulador. Además de actuar como enzima, la proteína hace las veces de represor. Se une a un sitio localizado entre el gen de la transposasa y su propio gen, evitando con ello la expresión de ambos genes.

No todos los transposones bacterianos son tan pequeños ni tan simples como el Tn3. Algunos son elementos



7. TRANSPOSICION DE *Mp*, elemento genéticamente idéntico a *Ac*, que ocurre (*a*) tras la replicación del elemento, pero antes de que el cromosoma donde reside acabe su duplicación. Una de las dos copias hijas del elemento se desplaza, generalmente a otra posición del mismo cromosoma. Si el sitio receptor no se ha replicado todavía (*b*), un cromosoma hijo llevará dos copias de *Mp* y el otro una sola. Si el sitio receptor se ha replicado ya (*c*), un cromosoma hijo no llevará el elemento *Mp* y el otro dos copias.



8. TRANSPOSON BACTERIANO Tn3. Mide unos 5000 nucleótidos de longitud y encierra tres genes. De éstos, dos codifican enzimas que catalizan la transposición: una "transposasa" y una "resolvasa"; el tercer gen codifica la enzima beta-lactamasa, que inactiva al antibiótico ampicilina. La resolvasa actúa de represor y de enzima. Se une a un sitio de una región no codificadora (*gris*), ubicada entre los genes de la transposasa y de la resolvasa, e impide la expresión de ambos. El represor cubre los sitios de iniciación donde comienza la transcripción de los genes en ARN; los sitios de iniciación son adyacentes porque los genes se transcriben en direcciones opuestas (*flechas de color*). En los extremos aparecen "repeticiones invertidas" de 38 nucleótidos: segmentos de ADN cuyas secuencias son simétricas y se leen del mismo modo, pero en direcciones opuestas y filamentos opuestos. Son señales de reconocimiento.

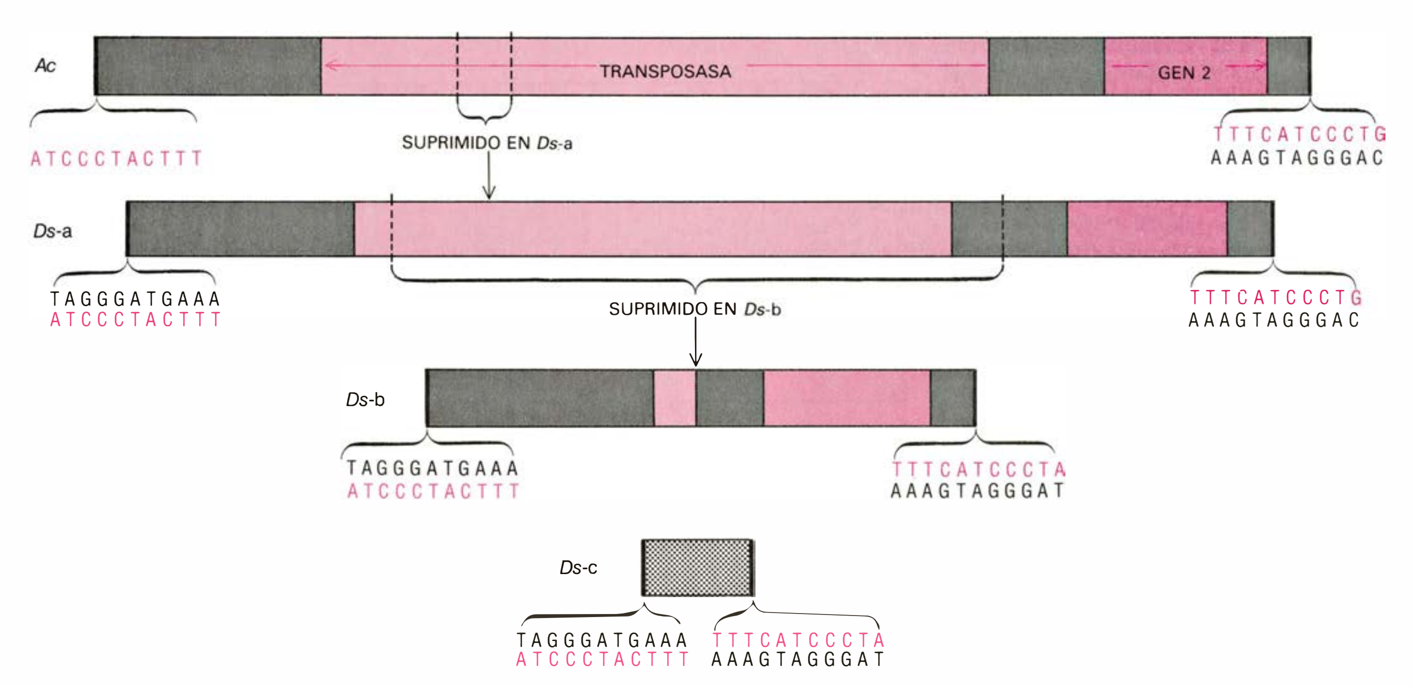
compuestos, provistos en ambos extremos de una copia de un pequeño elemento transponible. Se han aislado secuencias de ADN parecidas a transposones bacterianos en cierto número de organismos superiores, tales como levaduras, la mosca del vinagre *Drosophila melanogaster* y gusanos nematódos. Sin embargo, dada la abundancia de las copias de tales secuencias, y dado que se mueve con poca frecuencia, ha sido difícil demostrar que se transponen. Allan C. Spradling y Gerald M. Rubin, de la Institución Carnegie de Washington, demostraron que si un elemento *P* de *Drosophila* se inyectaba en un embrión de esa mosca con un plásmido promovía su propia transposición a un cromosoma de *Drosophila*, así como la de un elemento *P* "deficiente" portador de un marcador genético. Sus resultados evidencian que los elementos *P* tienen las propiedades de los transposones bacterianos y que también se parecen a los elementos *Ac* y *Ds* del maíz.

Aislamiento de *Ac*

Al emprender con mis colegas el aislamiento de *Ac* y *Ds* sabíamos que ambos elementos producían mutaciones por inserción en diversos genes. Supusimos que, si lográbamos aislar un gen con una mutación por inserción de un *Ac* o de un *Ds*, podríamos identificar qué parte del ADN aislado constituía el elemento transponible. Decidimos aislar el locus *waxi*, que codifica una enzima necesaria para la síntesis de la amilosa, uno de los dos tipos de almidón de la semilla del maíz. El endospermo de una semilla con un gen funcional *waxi* contiene amilosa y es bastante traslúcido. El endospermo de una semilla mutante que carece de amilosa tiene apariencia algo más opaca [*véanse las semillas* d-f *de la figura 3*].

Aislamos primero el ARN mensajero de semillas normales. (El ADN de un gen no se traduce directamente a proteína. Primero, una cadena de doble hélice de ADN –la cadena codificadora– se transcribe a una cadena complementaria de ARN, un ácico nucleico similar, que luego se traduce a proteína.) Realizamos la transcripción inversa de las moléculas de ARN de la semilla a copias de ADN complementario (ADNc), clonamos el ADNc en un plásmido e identificamos el clon que contenía el ADNc procedente del ARN mensajero codificado por el locus *waxi*. Ese ADNc era, en realidad, un gen *waxi* artificial, que serviría de sonda para encontrar el verdadero gen *waxi*, primero en plantas normales y después en plantas con mutaciones *Ac* y *Ds*.

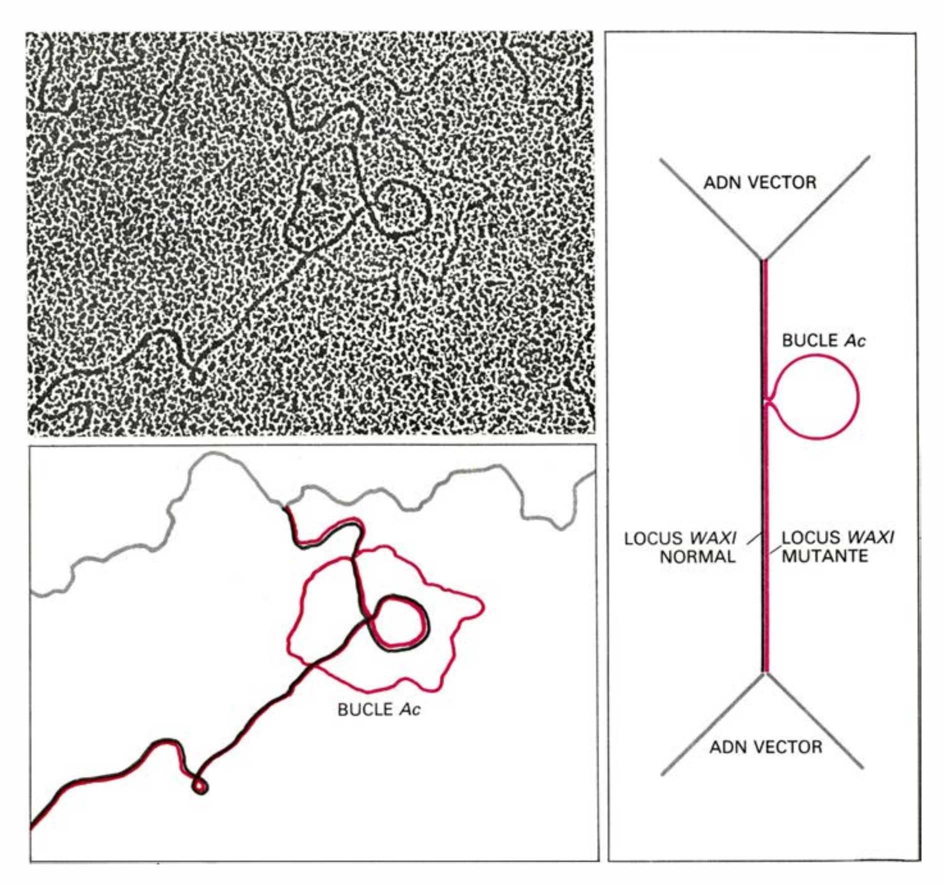
Para localizar el gen *waxi* extrajimos el ADN de plantas normales, lo cortamos en fragmentos breves e insertamos los fragmentos en el ADN de un vector



9. ELEMENTO *Ac* DEL MAIZ y varios elementos *Ds*, todos los cuales se han aislado ya. El elemento *Ac* tiene una longitud de unos 4500 nucleótidos; comprende dos secuencias codificadoras, o genes (*color*), tres secuencias no codificadoras (*gris*) y segmentos repetidos invertidos terminales, imperfectos (*negro*). El primer elemento *Ds* aislado, *Ds*-a, es casi idéntico a *Ac*, salvo en que un segmento de 194 nucleótidos (*líneas de trazos*) ha desaparecido por deleción del gen *Ac*, que es mayor. Esa pérdida basta para incapacitar al gen de la transposasa; y es responsable de la incapacidad de *Ds* para transponerse por sí mismo. En *Ds*-b ha desaparecido por deleción un segmento central, mucho más largo, de *Ac*. El gen *Ds*-c es muy corto; sólo conserva los segmentos repetidos terminales de *Ac*. Parece que cualquier segmento de ADN flanqueado por esas secuencias terminales puede actuar como elemento *Ds*.

clonante, el virus bacteriano llamado lambda. Cuando se añaden esas partículas víricas a un cultivo de células bacterianas, infectan las células y proliferan en gran número de clones víricos, cada uno de los cuales desciende de una sola partícula. Un clon puede contener millones de partículas víricas, es decir, millones de copias del fragmento de ADN de maíz insertado. Para encontrar el gen *waxi* entre esos fragmentos rastreamos los fragmentos de ADN clonado de maíz con nuestra sonda de ADNc marcado con una señal radiactiva. La sonda marcada se unía a los clones de ADN de maíz que contenían el gen *waxi*, identificándolos.

Aislado el gen *waxi* normal, repetimos el procedimiento con ADN de una estirpe portadora de una mutación inestable *Ac* en el locus *waxi*. (En esa estirpe, el endospermo es variegado para el tipo de almidón. Presenta sectores traslúcidos y sectores opacos.) Al comparar el gen normal con el mutante encontramos que éste era más largo. Poseía una pequeña secuencia de la que carecía el gen normal: la inserción *Ac*. Logramos visualizar en una microfotografía electrónica el segmento particular de ADN que correspondía al elemento *Ac* mediante el proceso conocido por heterodúplex, en el cual se combinan fragmentos de ADN que llevan el gen normal con fragmentos portadores del gen que contiene *Ac*. Al efecto se fomentó el apareamiento de cadenas simples de los dos fragmentos, todavía en su ADN vector. En la mayor parte de su longitud, las dos cadenas del ADN del maíz eran complementarias, razón por la cual en esas porciones se ensamblaban perfectamente para formar una molécula de ADN de doble cadena. Sin embargo, el heterodúplex tenía un bucle de cadena simple: una breve secuencia de una de las cadenas carecía de secuencia complementaria en la otra cadena. Puesto que los fragmentos normales *waxi* procedían de una estirpe revertida, derivada del abandono del locus por parte de *Ac*, no cabía duda de que el bucle era el propio elemento *Ac*.



10. ANALISIS POR HETERODUPLEX. Localiza aquí un elemento *Ac* inserto en el locus *waxi* del maíz. Se incuban monofilamentos de ADN clonado que llevan, ya el locus *waxi* normal, ya un locus mutado por inserción de *Ac*. Las secuencias complementarias se reconocen y se aparean en un ADN bicatenario, que se distingue como una línea más gruesa en la microfotografía electrónica. Esta se interpreta en los diagramas adyacentes. En la mayor parte de su longitud el locus normal (*negro*) y el mutado (*color*) se aparean, pero la inserción *Ac* no encuentra secuencias complementarias y forma un bucle de cadena simple.

Estructura de *Ac* y *Ds*

Con el tiempo logramos aislar el elemento *Ac*, analizar su estructura y determinar la secuencia de sus nucleótidos, en colaboración con Robert F. Pohlman y Joachim W. Messing, de la Universidad de Minnesota. *Ac* tiene poco más de 4500 nucleótidos; no es, pues, mucho más largo que el más corto de los elementos con capacidad reconocida para transponerse autónomamente. Es ligeramente más breve que el transposón bacteriano Tn3 y, curiosamente, semejante a él en organización. El elemento tiene segmentos repetidos terminales e invertidos, la marca característica de los elementos transponibles. En *Ac* la repetición es ligeramente imperfecta: las secuencias repetidas no superan los 11 nucleótidos. En la mayoría de los elementos transponibles las repeticiones terminales son más largas (38 nucleótidos en el Tn3), pero según parece las señales verdaderamente importantes para la transcripción las proporciona un número bastante reducido de nucleótidos, no más de 20, localizados en los extremos del elemento.

El elemento *Ac* se parece a un Tn3 sin su gen viajero de la beta-lactamasa. *Ac* tiene dos genes, uno largo y otro corto. Las proteínas que codifican son de un tamaño parecido al de la transposasa y la resolvasa del Tn3. Como ocurre en el transposón bacteriano, los dos genes están orientados de tal modo que se transcriben en direcciones opuestas, alejándose mutuamente. Se sabe que en Tn3 esa orientación permite el control coordinado de ambos genes por medio de la resolvasa, que actúa de represor al unirse a un solo sitio ubicado entre ambos. La asombrosa semejanza de la orientación génica en *Ac* sugiere que sus genes también podrían regularse desde un sitio común.

Existen pruebas de que el mayor de los dos genes de *Ac* codifica una proteína análoga a la transposasa de Tn3. La evidencia procede del análisis de un elemento *Ds* que aislamos. Como ya he mencionado, los datos genéticos sugerían que los elementos *Ds* guardaban estrecha relación con los *Ac*. No sólo se movilizaban los elementos *Ds* en presencia de *Ac*, sino que las mutaciones por inserción de *Ac* producían también espontáneamente mutaciones que actuaban como si se hubiera reemplazado un *Ac* por un elemento *Ds*. Aislamos el elemento *Ds* precisamente de uno de tales derivados de la mutación original por *Ac* del gen *waxi*. La inserción *Ds* se localizaba precisamente donde estaba la *Ac* en la estirpe parental. Se comprobó que la inserción era casi idéntica al elemento *Ac*, pero no completamente. Se parecía a un *Ac* mutante: faltaba una pequeña parte de la secuencia *Ac*. La longitud de la secuencia per-

dida era solamente de 194 nucleótidos. La deleción afectaba a la porción del gen *Ac* que codifica para la mayor de las dos proteínas. Un elemento *Ds* solamente podía transponerse en presencia de un *Ac*, de lo que dedujimos que la proteína cuya secuencia se había alterado por la deleción en el elemento *Ds* era una proteína implicada en la transposición: una transposasa.

No se sabe con certeza si todos los elementos *Ds* son meros *Ac* deficientes, o incompletos. El muestreo de los elementos *Ds* aislados hasta ahora sugiere que su estructura puede diferir mucho de la de los *Ac*. El segundo *Ds* que aislamos era la mitad de largo que *Ac*; estaba formado por los dos extremos de éste. En el laboratorio de Peter Starlinger, de la Universidad de Colonia, se aisló un elemento parecido. Consta de dos copias de un elemento semejante al menor de nuestros *Ds*, con una copia insertada, en orientación inversa, en la mitad de la otra copia. En el laboratorio de W. J. Peacock, de la Commonwealth Scientific and Industrial Research Organization de Canberra, se ha aislado un elemento muy distinto. Su longitud es diez veces menor que la del elemento *Ac*; se parece a éste en que tiene virtualmente los mismos segmentos repetidos terminales invertidos. Apunta ello la posibilidad de que cualquier secuencia flanqueada por esas dos cortas secuencias terminales pueda actuar de elemento *Ds*. Quizá toda la información que necesitan las enzimas de transposición para el reconocimiento, corte y transposición de una secuencia la den los segmentos repetidos invertidos. Si eso es así, cada vez que se presente en el genoma un elemento *Ac* para aportar las enzimas, cualquier segmento de ADN con secuencias terminales que reconozcan las enzimas será potencialmente capaz de desplazarse.

Familias de elementos

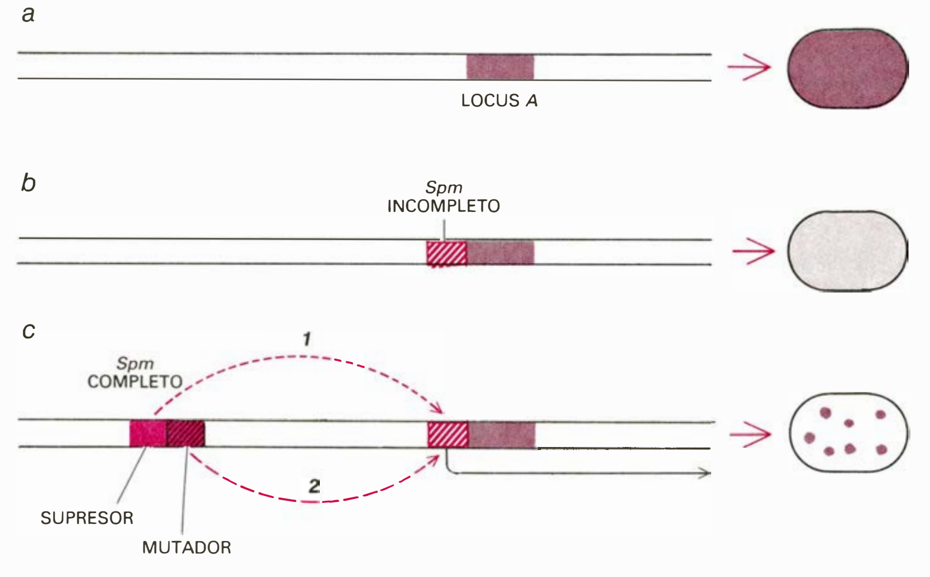
Se pensó en un principio que *Ac* constituía un solo locus. Estudios genéticos posteriores sugirieron que en un mismo genoma podían actuar simultáneamente más de un *Ac*, si bien el número de tales elementos sería pequeño. Sin embargo, los estudios genéticos no pueden detectar más que los elementos activos que tengan algún efecto observable sobre el organismo; las sondas moleculares para la detección de *Ac* de que dispusimos para el aislamiento del elemento podían detectar cualquier secuencia relacionada con *Ac*, fuera o no activa. El genoma parece salpicado de tales secuencias. Muchas de ellas se asemejan a los extremos del elemento *Ac*. Aunque las estimaciones no son muy exactas, parece que hay por lo menos 40 secuencias (y podría haber más de 100) similares a los extremos del elemento *Ac* en los genomas de todas las plantas que hemos examinado.

Pero los hallazgos moleculares no contradicen, en su conjunto, los hallazgos genéticos. Si nos preguntamos cuántos elementos son estructuralmente idénticos a *Ac*, la respuesta es que muy pocos. Hay razones para pensar que el elemento *Ac* que se detecta genéticamente tiene una estructura única. Aunque hay muchos trozos del elemento distribuidos por el genoma, se dan pocos elementos *Ac* completos. Muchos de los pedazos, aunque no sabemos cuántos, quizá se transpongan en presencia de un *Ac* funcional. En otras palabras, los elementos *Ac* y *Ds* constituyen una familia de elementos emparentados.

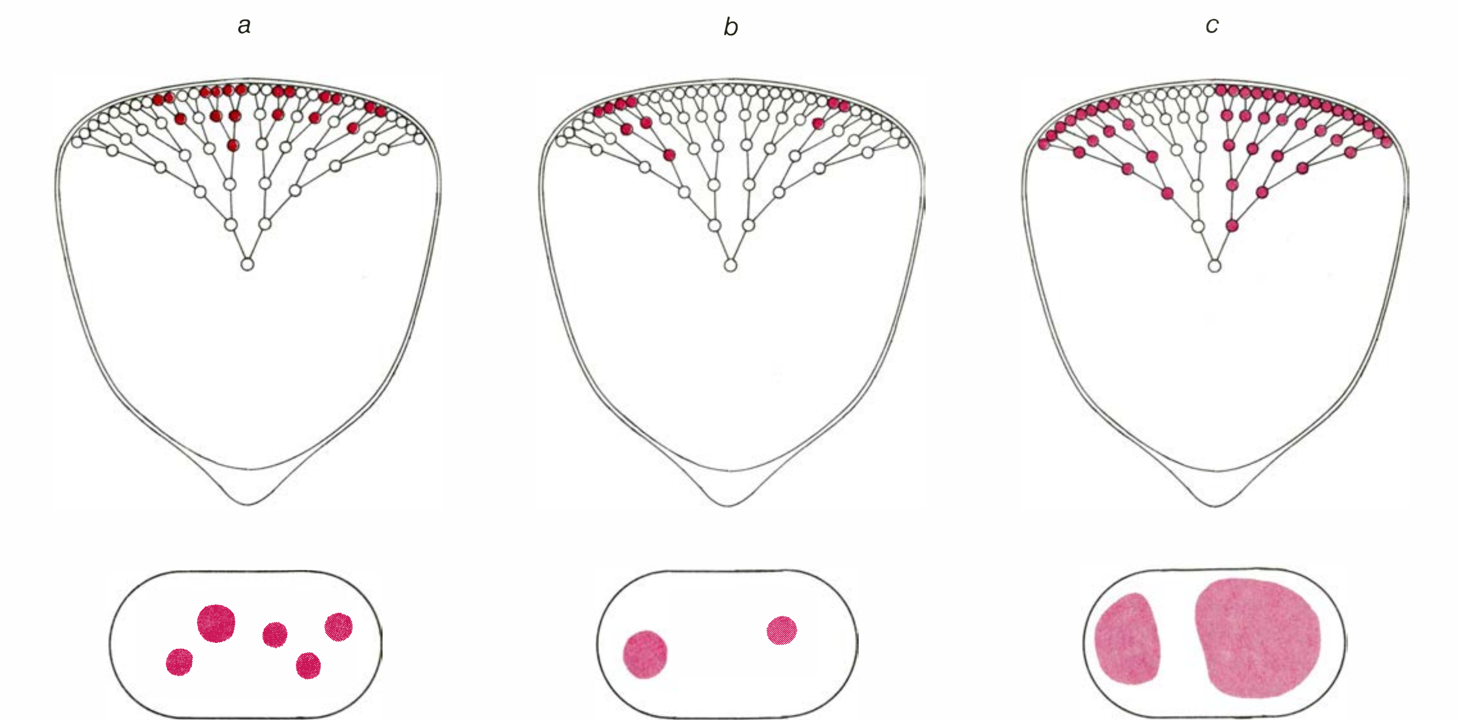
Los elementos *Ds* pueden considerarse una colección de secuencias más o menos relacionadas con *Ac*, portadoras todas ellas, por lo menos, de las secuencias críticas terminales que reconocen y transportan las enzimas de transposición de *Ac*. Por otra parte, los elementos *Ac* probablemente no tengan estructuras muy distintas, pues deben mantener su capacidad de codificar y expresar las enzimas de transposición. Se han aislado hasta ahora tres elementos *Ac* y parece que son casi idénticos. Sin embargo, no sorprendería encontrar uno que haya capturado un gen viajero, como parece haber hecho el Tn3 bacteriano en algún momento de su historia evolutiva.

Elementos *Spm*

Probablemente haya en el maíz muchas familias de elementos semejantes a la familia *Ac-Ds*. Es característica de los elementos transponibles que ciertas enzimas de transposición de elementos reconozcan solamente sus propias señales de transposición. Dicha especificidad permite dividir los elementos en familias, basándose en los experimentos genéticos. Así, puede introducirse un elemento transponible recién identificado en una planta que contenga una mutación producida por una inserción *Ds*. Si la mutación revierte, el nuevo elemento identificado es similar o idéntico a *Ac*; si la mutación permanece estable, el elemento pertenece a una familia diferente. Aplicando tal criterio se han identificado ya por lo menos seis familias de elementos en el maíz; quiza sólo correspondan a una pequeña fracción de las familias que realmente tiene esa especie. Los indios de América Central y del Sur, que fueron los primeros en cultivar el maíz, apreciaban las estructuras de pigmentación producidas por las mutaciones provocadas por elementos transponibles; se dispone por ello de muchas razas de maíz que muestran variegación. Solamente unas pocas se han recolectado y analizado genéticamente, por lo que no sabemos si sus mutaciones son producto de ele-



11. SUPRESOR-MUTADOR (*Spm*) es una familia de elementos transponibles que comprende elementos *Spm* "deficientes", o incompletos, análogos a *Ds*, y *Spm* completos, análogos a *Ac*. Aquí un *Spm* incompleto se ha insertado en el locus *A* (*a*), necesario para la síntesis de un pigmento púrpura de la aleurona, tras lo cual el gen pierde parte de su actividad (*b*); la semilla muestra, por tanto, una pigmentación pálida. La presencia de un *Spm* completo ejerce dos efectos (*c*). Su función supresora inactiva completamente al gen (*1*), lo que determina la formación de una semilla incolora. Sin embargo, por su efecto mutador (*2*), una transposasa sintetizada por el *Spm* completo induce al *Spm* incompleto a abandonar el locus *A* en algunas células durante el desarrollo; ello determina la aparición de sectores de células pigmentadas de oscuro.



12. TIEMPO DE DESARROLLO Y FRECUENCIA de transposición difieren en mutaciones producidas por la inserción de diferentes elementos *Spm* incompletos. Si la transposición se produce en una fase avanzada del desarrollo de la semilla, los clones de células revertidas son pequeños; también las manchas pigmentadas (*a*). Si la transposición tiene lugar en ese mismo momento pero con frecuencia más baja, aparecen menos clones y, por tanto, menos manchas (*b*). Con una precoz transposición que restaure la función, los clones revertidos y las manchas de tejido pigmentado serán mayores (*c*).

mentos transponibles conocidos o de elementos todavía por identificar.

Sólo se han estudiado con el detalle de los elementos *Ac* y *Ds* los pertenecientes a la familia *Spm*. Esta familia de elementos la identificaron independientemente McClintock y Peter A. Peterson, de la Universidad estatal de Iowa, en la década de 1950. McClintock llamó elemento *supresor-mutador* (*Spm*) al miembro plenamente funcional de la semilla; Peterson lo denominó *Enhancer* (destacador). No se ha aislado aún, pero las mutaciones producidas por los elementos *Spm* sí se han estudiado con gran pormenor en su vertiente genética. Las mutaciones ilustran vívidamente el complejo de efectos de los elementos transponibles en los organismos superiores; proporcionan al biólogo molecular una rica fuente de material cuyo análisis quizás ilumine algún día los mecanismos que determinan cómo se expresan los genes en diferentes estadios del desarrollo y en tejidos específicos.

El comportamiento básico de los elementos *Spm* se parece al de los *Ac* y *Ds*. Un elemento *Spm* autónomo, lo mismo que un *Ac*, puede producir una mutación inestable insertándose en un gen, o en su vecindad. La inserción inactiva el gen; la posterior escisión del elemento en algunas células durante el desarrollo restaura la función del gen, dando a las semillas la apariencia variegada. La familia *Spm* comprende también elementos no autónomos análogos a los *Ds*. Tales *Spm* incompletos, o "deficientes" (McClintock no acuñó ningún término con el que designarlos) pueden transponerse solamente cuando está presente un *Spm* autónomo completo (o no deficiente). A pesar de esta semejanza básica entre *Spm* y *Ac*, se trata claramente de elementos diferentes, según el criterio que se ha descrito. Un elemento *Spm* no puede hacer que se desplace un elemento *Ds*, y un *Spm* incompleto no se desplaza en presencia de *Ac*.

Se aprecia otra interesante diferencia funcional entre las familias *Spm* y *Ac*. Los elementos *Spm* incompletos se insertan frecuentemente en un gen de suerte tal que el gen sólo se inactiva parcialmente. En tanto el genoma no incluya un *Spm* incompleto el gen continúa expresándose, aunque con menos intensidad; si la mutación afecta a un gen necesario para la síntesis de un pigmento, la semilla aparece con pigmentación pálida, en lugar de incolora [*véase la semilla* g *de la figura 3*].

La situación cambia por entero en presencia de un *Spm* completo. El gen de pigmentación queda "suprimido", esto es, inactivado del todo; la semilla aparece entonces incolora. Sin embargo, el *Spm* completo aporta una transposasa, por lo que el elemento incompleto se escinde en algunas células en el curso del desarrollo, produciendo manchas profundamente pigmentadas, formadas por células que han revertido y en las que el gen funciona normalmente [*véanse las semillas* h-j *de la figura 3*]. McClintock denominó *supresor-mutador* a tal elemento, porque ejerce dos efectos distinguibles sobre un gen que transporte un elemento incompleto: supresión y retromutación, o reversión.

Mutaciones de *Spm*

McClintock estudió dos categorías de cambios hereditarios en los propios elementos *Spm*: mutaciones que afectan a la transposición de elementos incompletos y mutaciones que afectan a las funciones de *Spm* completos. Todas las mutaciones de la primera categoría producen una estructura de manchas pigmentadas, pero las estructuras de variegación pueden ser muy diferentes. Cada mancha corresponde a un clon de células derivadas de una sola en la que el *Spm* incompleto se escindió del gen. Si la transposición se produce en un estadio avanzado del desarrollo de la semilla, cuando solamente faltan pocas divisiones celulares, los clones de las células que hayan revertido (las manchas) serán pequeños. Que las manchas sean más o menos numerosas dependerá de la frecuencia con la que el *Spm* incompleto se transponga; cada mutación goza de una frecuencia característica. Si la transposición que restaura la función génica tiene lugar al poco de iniciarse el desarrollo de las semillas,

los clones de células revertidas son grandes [*véanse las semillas* h-j *de la figura 3*]. Por tanto, mutaciones que afecten al mismo locus, producidas por lo que a nivel genético parece ser un mismo *Spm*, pueden diferir tanto en la frecuencia como en el momento del desarrollo en el que el elemento se escinda del locus.

¿Qué es lo que hace tan distintas a esas mutaciones? ¿Cómo "sabe" el elemento cuándo y con qué frecuencia transponerse? ¿Contiene cada elemento en su secuencia de nucleótidos una señal genética que actúa a modo de reloj del desarrollo, o es el elemento, en cierto sentido, un cautivo del programa de desarrollo del gen en el cual se inserta? ¿Acaso, cuando el elemento se inserta en puntos cercanos, o del propio gen, queda accesible a las enzimas de transposición en distintas fases del desarrollo de la semilla? No puede aún responderse a esas preguntas, pero merece la pena planteárselas. A medida que se aíslen genes mutantes y se estudie la estructura y localización de los elementos *Spm* incompletos que contienen, se recabará más información acerca de las señales genéticas que determinan las secuencias temporales del desarrollo.

Algunas mutaciones que afectan a las funciones del elemento *Spm* completo guardan un estrecho parecido con las mutaciones del elemento incompleto que acabo de describir. Influyen sobre la frecuencia y momento de la transposición, en algunos casos incapacitando al elemento para promover su propia transposición o la de algún elemento incompleto; y, sin embargo, manteniendo intacta su función supresora. Quizá las mutaciones más intrigantes sean las que afectan a la expresión global de los elementos completos. En algún caso, el elemento parece ponerse en funcionamiento y dejar de funcionar repetidamente durante el desarrollo, esto es, atraviesa un ciclo de estados activos e inactivos.

Se demuestra mejor el comportamiento cíclico de este *Spm* por su efecto sobre una mutación producida por cierto *Spm* incompleto que es incapaz de transponerse incluso en presencia de un elemento completo, si bien responde a la función supresora del elemento. En ausencia de un *Spm* completo, las semillas que llevan esa mutación aparecen profundamente pigmentadas. En presencia de un *Spm* completo normal, las semillas carecen absolutamente de color, en respuesta a la función supresora del elemento. No presentan manchas, pues el elemento incompleto (deficiente) está retenido y no puede escindirse para restaurar la función génica normal.

Cuando se reemplaza el *Spm* completo (no deficiente) usual por la versión cíclica de ese elemento, las semillas muestran un aspecto distinto de su variegación. Se observan sectores pigmentados bastante grandes en la aleurona coloreada; dentro de esas grandes manchas hay, a su vez, pequeños sectores incoloros [*véase la semilla* k *de la figura 3*].

Esa estructura de variegación nada tiene que ver con la transposición del *Spm* incompleto, que no puede desplazarse. La estructura es producto del ciclo del *Spm* completo. Cuando el elemento está activo, expresa su función supresora y resulta incolora la porción correspondiente de la aleurona. Cuando el elemento está inactivo, la sustancia supresora no se sintetiza y el gen funciona, produciendo sectores pigmentados. Sin embargo, el elemento solamente se silencia de modo transitorio; al cabo de varias divisiones celulares más puede activarse de nuevo, dando lugar a pequeños sectores incoloros dentro de las manchas pigmentadas.

Otros *Spm* completos no muestran un ciclo de actividad e inactividad en la planta, sino que se activan en un momento del desarrollo o en una parte de la planta y suspenden su funcionamiento en otros momentos y en otros lugares. Un elemento puede mostrarse activo, por ejemplo, en la extremidad de la semilla, o en su base sólo. O puede estar activado en las mazorcas producidas en el tallo principal de la planta e inactivado en las producidas en los hijuelos, o tallos laterales.

Desarrollo y evolución

La ubicuidad de los elementos transponibles y la seguridad creciente de que se trata de agentes destacados de inestabilidad genética nos fuerzan a considerar su importancia en la coreografía genética que subyace al desarrollo y su papel en el proceso de cambio genético a largo plazo llamado evolución. McClintock denominó "elementos controladores" a los elementos genéticos que estudió; comprendió, en efecto, que en cierto sentido inhibían reversiblemente la expresión génica y que el tipo de escisión, determinado genética-

a
Spm INCOMPLETO PARA TRANSPOSICION
LOCUS *A-2*
b
Spm COMPLETO
SUPRESOR
MUTADOR
c
Spm CICLICO COMPLETO
Spm ACTIVADO
Spm INACTIVADO
ACTIVO
INACTIVO
ACTIVO

13. UN *Spm* CICLICO atraviesa alternativamente fases de actividad e inactividad. Se manifiesta dicho efecto cuando el *Spm* incompleto situado en un locus que determina una enzima de pigmentación no puede transponerse. En ausencia de un *Spm* funcional, el *Spm* incompleto, incapaz de transponerse, apenas influye en la pigmentación (*a*). En presencia de un *Spm* ordinario (*b*), su función supresora inactiva el gen, pero su función mutadora no se manifiesta, porque el elemento no se desplaza; la semilla no tiene color. Si el *Spm* es cíclico (*c*), la función supresora se activa y suprime alternativamente, de tal modo que aparecen sectores pigmentados (*Spm* inactivados) con sectores incoloros en su interior (*Spm* activados).

mente, que se da en células somáticas (las células no germinales) era responsable de una "estructura" muy característica de expresión génica del organismo considerado en su conjunto. Además, las propiedades de los elementos y sus interacciones con genes originaban múltiples y variadas estructuras de expresión génica. Ello sugirió a McClintock la posibilidad de que los elementos controladores fueran ejemplos de mecanismos normales de control que actuaran de manera inapropiada.

La idea de que los elementos transponibles fueran importantes mecanismos que controlan la expresión génica en el desarrollo no ha recibido apoyo experimental directo. No cabe duda, sin embargo, de que diversos genes, como los que determinan la síntesis de anticuerpos en animales superiores, se reestructuran en el curso de su programa normal de desarrollo. Además, una expresión génica de estructura muy parecida a la que se observa en el maíz se da también en otras plantas y animales. No es aventurado sospechar que los mecanismos genéticos básicos responsables de tales estructuras se asemejan a los descritos en el maíz.

Quizá los elementos transponibles desempeñen una labor aún más trascendental en la evolución. No nos cabe sino especular sobre ese hipotético papel, pero sus propiedades parecen hacerlos agentes aptos para modificar, además de la expresión de los genes, la estructura de genes y genomas. En el maíz y otros organismos se sabe que los elementos transponibles se activan cuando el genoma está sometido a tensiones y se producen roturas cromosómicas. Ignoramos los mecanismos del proceso; ahora bien, en cuanto los elementos se activan, pueden promover muchos tipos de mutaciones y reorganizaciones cromosómicas. Es como si los elementos transponibles pudieran amplificar una pequeña alteración, transformándola en un terremoto genético. Quizá tal turbulencia genética sea una fuente importante de variabilidad genética, materia prima de la cual la selección natural recoge lo que sea útil para las especies. Por otra parte, según indican pruebas recientes, además de inactivar los genes, los elementos transponibles pueden ponerlos en marcha o potenciar su expresión. Hay razones para sospechar que pueden también reprogramar los genes de modos más sutiles, cambiando el momento y lugar de su activación. Ello constituye, de hecho, la materia prima del remodelado y la reconstrucción, en suma, de la evolución orgánica.

Preferencia atómica entre izquierda y derecha

Las versiones levógira y dextrógira de un experimento atómico no siempre dan lugar a efectos especulares simétricos. La asimetría está calculada por la fuerza débil entre los electrones y el núcleo

Marie-Anne Bouchiat y Lionel Pottier

La idea de simetría es fundamental para entender la naturaleza. Un sistema físico que exhibe un alto grado de simetría es el átomo. En los átomos más sencillos, los electrones que orbitan alrededor del núcleo están distribuidos en capas esféricas, obviamente simétricas en todas las orientaciones posibles. En átomos más complejos, la distribución de los electrones no es esférica; pero incluso aquí se puede identificar algún plano o eje de simetría. Para el físico existe una simetría en el átomo todavía más profunda: hasta años recientes, el átomo se contemplaba como un sistema gobernado tan sólo por la fuerza electromagnética, fuerza que es totalmente indiferente a la distinción entre izquierda y derecha. De modo más preciso, supongamos que cualquier configuración física y una segunda configuración exactamente como la imagen especular de la primera son sometidas las dos a la fuerza electromagnética. Ocurre, entonces, que el efecto de la fuerza sobre la segunda configuración es idéntico a la imagen especular del efecto de la fuerza sobre la primera configuración. Esta propiedad importante de la fuerza electromagnética se llama conservación de la paridad.

En la última década se ha realizado un notable esfuerzo experimental para examinar la conservación de la paridad en el átomo. Tales ensayos no están motivados por ninguna sospecha de que la fuerza electromagnética viola o, en otras palabras, no conserva la paridad. Por el contrario, la influencia de la fuerza electromagnética sobre las interacciones atómicas es tan poderosa que cualesquiera desviaciones observadas de la conservación exacta de la paridad se espera que sean extraordinariamente pequeñas. Ahora bien, el átomo no es un sistema puramente electromagnético. Se cree que tres fuerzas fundamentales, además de la fuerza electromagnética, contribuyen a los movimientos de los constituyentes elementales del átomo. Las tres fuerzas son: la gravedad, la fuerza fuerte, o nuclear, y la fuerza débil. De las tres, la gravedad es despreciable a escala atómica; la fuerza fuerte no afecta a los electrones, cuyos movimientos determinan las propiedades de todas las interacciones atómicas. Pero, de acuerdo con los cálculos teóricos, la fuerza débil entre cada electrón y el núcleo atómico debería causar pequeñas perturbaciones en las órbitas de los electrones, con tal de que éstos discurran suficientemente cercanos al núcleo para "sentir" los efectos de la fuerza débil.

Las perturbaciones causadas por la fuerza débil deberían dar lugar a una pequeña violación de la paridad en el átomo. Los efectos de la fuerza débil en el átomo pueden examinarse, pues, en un ensayo experimental de un modo que es conceptualmente directo, aunque, en ningún caso, tan sencillo en la práctica: a la configuración de los átomos de un gas se la dota de una lateralidad preferente, por ejemplo izquierda, y se mide alguna propiedad adecuada de los átomos. La configuración se reemplaza entonces por otra que coincida con la imagen especular de la primera configuración; la segunda configuración ha de tener, por tanto, una lateralidad preferente opuesta a la de la primera. Se mide, entonces, la misma propiedad atómica para la segunda configuración. Si la propiedad medida en el segundo ensayo no es la imagen especular exacta del primer resultado, se viola la paridad.

Para finales de la década de los setenta, habían fracasado ya varios experimentos que pretendían detectar tales violaciones de paridad. Sus resultados empezaron a considerarse un débil eslabón en una cadena de razonamiento teórico que, por otro lado, tenía pleno éxito. Este importante eslabón ha quedado ahora forjado. Varios grupos de investigadores, incluido el nuestro de la École Normale Supérieure de París, han confirmado la existencia de violaciones de paridad en algunos sistemas atómicos. Esos resultados completan los éxitos espectaculares de la nueva teoría de interacciones débiles a la hora de predecir fenómenos a energías extremadamente altas, tales como el descubrimiento, este último año, de las partículas W^+, W^- y Z^o en el Laboratorio Europeo para Física de Partículas (CERN), en Ginebra. Los experimentos en física atómica se llevan a cabo a energías mucho más bajas que las generadas en los aceleradores, por lo que examinan la acción de la fuerza débil sobre una distancia mucho mayor de la que lo hacen los aceleradores. Además, proporcionan información sobre la magnitud de la fuerza débil entre los constituyentes fundamentales de la materia, como quarks y electrones, no deducida aún a partir de los ensayos de alta energía.

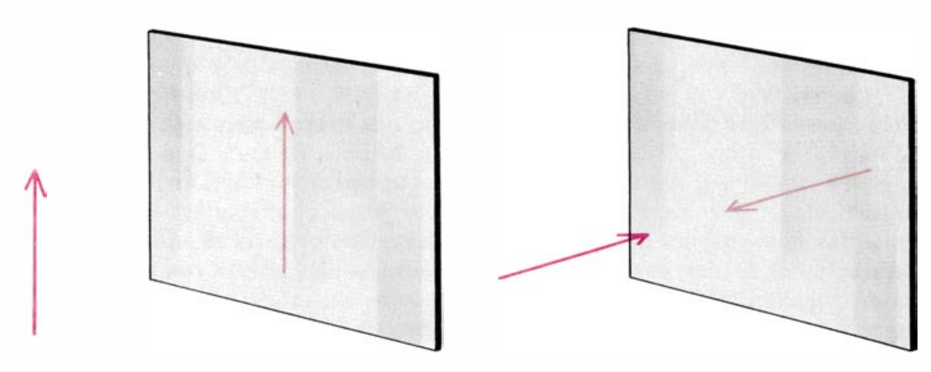
La fuerza débil es asociada comúnmente con la desintegración beta, paso importante en cierto mecanismo para la conversión termonuclear en el Sol de hidrógeno en helio. Un neutrón se desintegra en un protón en el interior de un núcleo atómico; la reacción viene acompañada por la emisión de un electrón (el "rayo beta") y una partícula llamada antineutrino. Se sabe desde hace más de 25 años que la paridad no se conserva en esta reacción. El electrón sale siempre del núcleo con un espín en el mismo sentido en que los dedos de la mano izquierda se enroscarían si el pulgar izquierdo señalara la

dirección del movimiento del electrón. El antineutrino se emite siempre con el espín en sentido dextrógiro. La imagen especular del proceso sería una reacción en que el electrón "beta" llevara espín en el sentido dextrógiro y el antineutrino en sentido levógiro; tal reacción no se observa nunca en el mundo real.

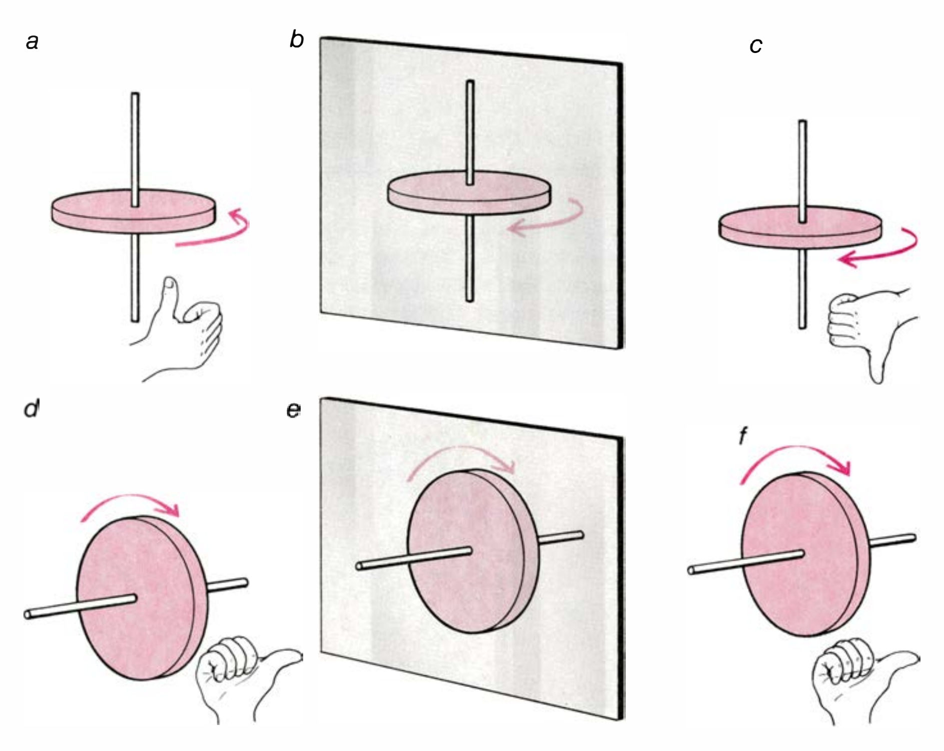
La violación de la paridad en la desintegración beta va acompañada por un cambio en la carga eléctrica del neutrón que se desintegra. En cuanto se catalogaron otras reacciones débiles, se hizo evidente que se acompañaban siempre de un cambio en la carga y, por ello, en la identidad de la partícula que reacciona. Así, un neutrino puede convertirse (por la adición de energía) en un electrón, pero en el proceso ha de adquirir una unidad de carga eléctrica negativa. Esas transformaciones las producen las partículas que median la fuerza débil; las tales pertenecen a la familia de los llamados bosones de aforo.

Los bosones de aforo que median la fuerza débil son análogos al fotón, el bosón de aforo que media la fuerza electromagnética. Sin embargo, de modo distinto al fotón, eléctricamente neutro, los bosones de aforo débiles que median la desintegración beta (y todas las demás interacciones débiles conocidas antes de 1970) transportan una unidad de carga eléctrica. El bosón de aforo W^+ lleva una unidad de carga positiva, y el bosón de aforo W^-, una unidad de carga negativa. En consecuencia, se dio por sentado que ni la fuerza débil ni su violación asociada de la paridad importaban mucho en la física del átomo estable. Si la fuerza débil efectuaba siempre una transformación de carga, no podría afectar los movimientos de los electrones atómicos sin alterar su carga, minando con ello la estabilidad del átomo.

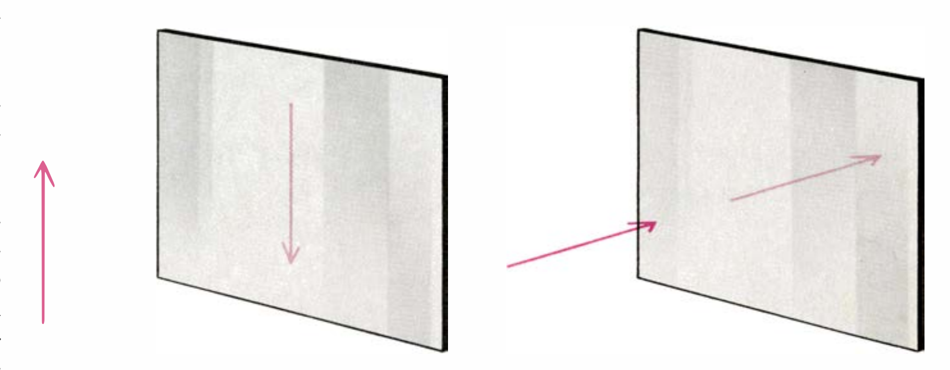
La comprensión teórica de las interacciones débiles resultó insatisfactoria, desde el punto de vista matemático, hasta finales de los años sesenta, cuando se propuso una nueva teoría. Sheldon Lee Glashow y Steven Weinberg, ambos, entonces, en la Universidad de Harvard, y Abdus Salam, del Centro Internacional de Física Teórica de Trieste, sugirieron, independientemente, que las fuerzas débil y electromagnética podían interpretarse como manifestaciones diferentes de una sola fuerza subyacente: la fuerza electrodébil. La teoría electrodébil unificada recibió más tarde una base matemática sólida con los trabajos de Gerard't Hooft, de la Universidad de Utrecht.



1. SIMETRIA Y ASIMETRIA en un sistema físico, el átomo por ejemplo. Pueden definirse considerando los efectos de reflexión especular sobre objetos tales como el vector ordinario, o flecha. Un espejo mantiene la dirección manifiesta de la flecha cuando aquél y la flecha se hallan paralelos (*dibujo de la izquierda*) e invierte, por contra, la dirección cuando se coloca el espejo perpendicular a la flecha (*derecha*).



2. SE LLAMA REGLA DE LA MANO DERECHA a un método convencional para designar adecuadamente el sentido de rotación de un objeto que gira sobre sí mismo. La regla establece que el sentido de rotación queda dado por una flecha alineada a lo largo del eje de rotación y apuntando en la dirección en que señalaría el pulgar derecho si los dedos de su mano se enroscaran en la dirección de giro (*a,d*). Cuando el plano de un disco en rotación es perpendicular a un espejo, la imagen del disco en el espejo gira en sentido opuesto al del disco real (*b*). Si el plano del disco es paralelo al espejo, los dos, el disco y su imagen, giran en el mismo sentido (*e*). La mano derecha no aparece reflejada en los espejos porque, de hacerlo así, daría a las imágenes especulares de los discos en rotación un sentido que contradice la regla de la mano derecha. Cuando un disco real gira en el mismo sentido que lo hace el disco de la imagen especular, la regla de la mano derecha nos descubrirá el sentido de la rotación sin ambigüedad (*c* y *f*).



3. LOS VECTORES AXIALES, que son las flechas que representan el sentido de rotación de un disco en giro sobre sí mismo, se transforman por una reflexión especular en una imagen que es la opuesta a la imagen de un vector ordinario. Si el vector axial es paralelo al espejo, su dirección se invierte por el espejo (*izquierda*); si el vector axial es perpendicular al espejo, su dirección permanece invariante (*derecha*).

De acuerdo con la teoría, hay un tercer bosón de aforo, neutro, llamado Z^o, que media un nuevo tipo de fuerza débil. Las interacciones que involucran al bosón de aforo Z^o pueden violar también la paridad; ahora bien, como la partícula Z^o no transporta carga eléctrica, las interacciones no conducen a un cambio en la identidad de las partículas que interaccionan. Así, el intercambio del Z^o es análogo al intercambio de un fotón en una interacción electromagnética. Las interacciones débiles que implican el intercambio de partículas Z^o fueron observadas en experimentos con neutrinos realizados en el CERN, en 1973. En principio, la teoría electrodébil había propuesto, pues, un mecanismo, cuya existencia fue parcialmente confirmada por la experiencia, según el cual los electrones atómicos de un átomo estable podían sentirse perturbados por la fuerza débil.

En el átomo real, sin embargo, y al hilo de unas sencillas consideraciones teóricas se deduce que los efectos de la fuerza débil sobre los electrones que orbitan serían demasiado pequeños para llegar a detectarse. La teoría electrodébil predice una enorme masa para los tres bosones de aforo débiles: los bosones W^+ y W^- han de tener cada uno una masa equivalente a una energía de unos 80 GeV (mil millones de electronvolt) y el bosón Z^o será algo más pesado, alrededor de 90 GeV. Todas esas predicciones fueron confirmadas, el año pasado, con los descubrimientos experimentales del CERN. Se sabe, desde las investigaciones, en 1934, de Hideki Yukawa, de la Universidad de Kyoto, en Japón, que cuanto mayor es la masa de una partícula, más corto es el alcance de la fuerza que media. Por ejemplo, los fotones que median la fuerza electromagnética tienen masa nula; el alcance de la fuerza es infinita. Las grandes masas de los bosones de aforo débiles determinan que el alcance de la fuerza débil sea extremadamente corto, alrededor de 2×10^{-18} metros. Este alcance ha recibido confirmación experimental de ensayos que medían la dispersión del neutrino por otras partículas, un proceso gobernado sólo por la fuerza débil. El alcance divide por más de diez millones de veces el radio típico de un átomo.

4. VIOLACION DE LA PARIDAD en la desintegración beta. Se ha venido interpretando, desde hace más de 25 años, como efecto de la fuerza débil. Un neutrón se desintegra dando un protón en el núcleo atómico, y emitiéndose un electrón y un antineutrino. El electrón posee siempre espín en el mismo sentido en que los dedos de la mano izquierda se enroscarían si el pulgar izquierdo señalara la dirección del movimiento del electrón. El antineutrino posee espín en un sentido dextrógiro con relación a su dirección de movimiento. La imagen especular del proceso está prohibida por las leyes de la fuerza débil.

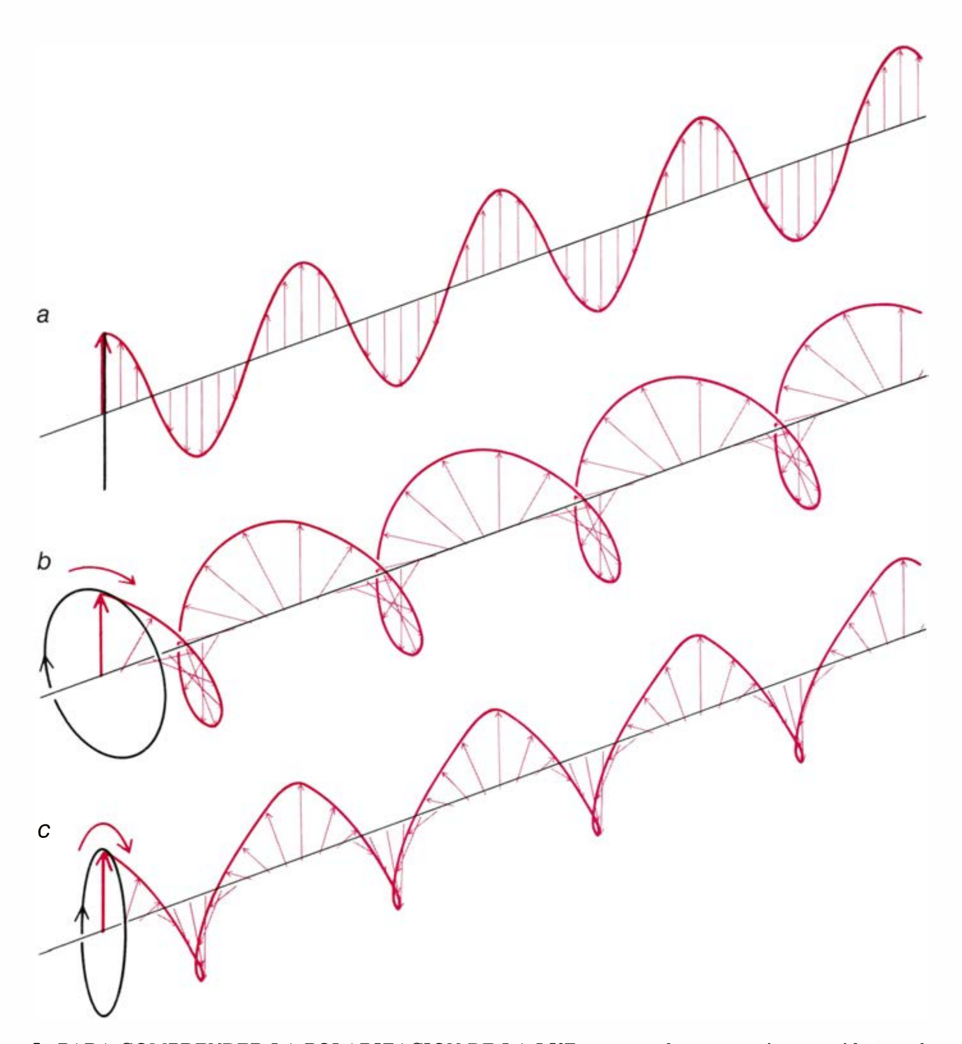
¿Cómo hemos de esperar que se detecten los efectos de la fuerza débil en el átomo? En la mayoría de las órbitas, el electrón atómico está virtualmente más allá del alcance de la fuerza débil; mas, si la órbita es altamente excéntrica o alargada, el electrón puede acercarse lo suficiente al núcleo para que la fuerza débil ejerza un influjo más poderoso. Como las propiedades espectroscópicas y químicas del átomo están determinadas, casi completamente, por los electrones más externos, uno de esos electrones tiene que penetrar por la nube de electrones más internos para que se pueda observar una violación de la paridad.

En la región interna de tal órbita, el movimiento del electrón penetrante no está afectado por los otros electrones atómicos, que, habitualmente, apantallan la carga positiva del núcleo. Antes bien, el electrón penetrante está sometido al efecto completo de todos los protones en el núcleo; la parte de la órbita electrónica que subyace a la nube de los otros electrones en el átomo se parece a la órbita del único electrón del átomo de hidrógeno. El núcleo de un átomo pesado está formado por muchas cargas positivas; el núcleo de hidrógeno tiene, por contra, una sola unidad de carga. (En cualquier átomo, la carga nuclear total es igual a Z, el número atómico.) De acuerdo con ello, la distancia entre el electrón y el núcleo en la región interna de la órbita electrónica es un factor Z menor que la distancia entre el electrón y el núcleo en el átomo de hidrógeno. Además, la fuerza débil es proporcional a la velocidad del electrón cerca del núcleo, lo cual aumenta por otro factor Z. En 1973, Claude C. Bouchiat, también de la École Normale Supérieure, y uno de nosotros (Marie-Anne Bouchiat) mostraron que la asimetría izquierda-derecha, la magnitud de la violación de paridad, esperada bajo tales circunstancias, aumentaba algo más rápidamente que el cubo de Z.

La amplificación cúbica aproximada de la violación de paridad implica que el efecto de la fuerza débil debería ser de 10^6 a 10^7 veces más fuerte en los átomos más pesados, en los que Z está entre 50 y 100, de lo que es en el hidrógeno. Pero hay una dificultad en los átomos pesados que debe contrapesarse contra su ventaja. Para aplicar los resultados de un experimento a la determinación de las constantes numéricas de la teoría electrodébil, se ha de dar una descripción matemática complicada del átomo. La dificultad del cálculo

aumenta con el número de electrones de la capa electrónica más externa. En una primera aproximación, las otras capas pueden abordarse como una estructura esférica bastante inerte.

Los físicos experimentales han encontrado diversos compromisos entre la simplicidad de cálculo y la magnitud del efecto al intentar medir la violación de paridad en sistemas atómicos. Se realizan ensayos con bismuto (número atómico 83), plomo (82), talio (81), cesio (55) e incluso hidrógeno (1). Hay ahora alrededor de una docena de experiencias en marcha o recientemente realizadas; podemos agruparlas en tres categorías principales.



5. PARA COMPRENDER LA POLARIZACION DE LA LUZ tracemos la trayectoria recorrida por el vector que representa el campo eléctrico de la luz, medido en cualquier punto fijado del espacio. El vector oscila en un plano perpendicular a la dirección de propagación del haz, con la frecuencia de la onda de luz. Si la punta del vector dibuja un segmento de línea recta en el plano de polarización, la luz queda polarizada linealmente; si la punta del vector traza una circunferencia, la luz está polarizada circularmente. En el caso más general, la punta del vector dibuja una elipse y la luz está polarizada elípticamente. Cuando la propagación del haz de luz se combina con el movimiento del vector de campo eléctrico, dicho vector traza una trayectoria en el espacio. Si la luz está polarizada linealmente, la trayectoria es una onda sinusoidal, sin lateralidad alguna (*a*). Si la luz está polarizada circularmente, la trayectoria es una hélice cuya lateralidad depende de la relación entre la rotación del vector y la dirección de propagación (*b*). En la hipótesis de la polarización elíptica, la trayectoria del vector es también levógira o dextrógira (*c*).

En la primera categoría se inscriben ensayos que miden la rotación óptica de la luz polarizada por un gas de átomos. Esta técnica fue propuesta, independientemente, por Patrick G. H. Sandars, de la Universidad de Oxford, I. B. Khriplovich, del Instituto de Física Nuclear de Novosibirsk, y David C. Soreide y Edward N. Fortson, de la Universidad de Washington. Para comprender la rotación óptica conviene atender a los modos en que un haz de luz puede polarizarse.

Existe un campo eléctrico asociado con todo punto a lo largo de un haz de luz. El campo eléctrico en cada punto es un vector, o flecha, que yace en el plano perpendicular a la dirección de propagación del haz y oscila rápidamente en ese plano. Cuando la punta de la flecha oscila a lo largo de una línea recta en el plano, la polarización de la luz es lineal; al propagarse el haz, la flecha oscilante traza una curva sinusoidal en el espacio. Cuando la punta de la flecha gira a lo largo de una circunferencia en el plano de polarización, la polarización del haz es circular; la traza de la flecha en el espacio es una hélice que se enrosca a través del espacio como un tornillo levógiro o dextrógiro, según sea el sentido de rotación de la flecha.

El tipo más general de polarización es la elíptica: la punta de la flecha describe una elipse en el plano de polarización. Tanto la polarización lineal como la circular son casos especiales de polarización elíptica. La polarización lineal constituye un ejemplo degenerado en el que el eje menor de la elipse es cero; la polarización circular es el caso extremo, en que los dos ejes de la elipse son iguales en longitud. La luz polarizada elíptica tiene una lateralidad definida, como la tiene la luz polarizada circularmente, porque el movimiento del campo eléctrico a lo largo de la elipse puede ser levógiro o dextrógiro, con relación a la dirección de la propagación del haz. La luz polarizada linealmente, sin embargo, carece de lateralidad; corresponde al caso intermedio, en que el movimiento de la flecha no es ni levógiro ni dextrógiro.

La rotación óptica es un efecto de un medio ópticamente activo sobre la luz polarizada linealmente. Cuando un haz de luz linealmente polarizado atraviesa tal medio, la polarización lineal del haz gira cierto ángulo y el haz emerge del medio polarizado linealmente en una dirección distinta. Dicha rotación se ha detectado en muchos compuestos moleculares, incluidas muchas sustancias biológicas. Hasta hace poco, sin embargo, no se había observado en un gas atómico. La razón estriba en que la rotación óptica molecular viene causada por la lateralidad de las partículas que componen el medio.

Es típico que los átomos de una molécula se organicen de modo tal que ninguna rotación o movimiento en el espacio transforme la molécula en una réplica de su imagen especular. La molécula de ADN, por ejemplo, se enrosca a lo largo de su eje central en un sentido preferente; la molécula no puede transformarse en su propia imagen especular, como no lo puede una mano derecha. Decimos que la molécula exhibe lateralidad. Si una molécula dextrógira exhibe una rotación óptica A, en el espejo se vería a la correspondiente molécula levógira exhibir la rotación óptica opuesta $-A$. Resulta que la rotación óptica causada por una molécula levógira real es precisamente

igual a $-A$; en otras palabras, se ajusta a la imagen especular de la rotación óptica de la molécula dextrógira. Así pues, en la rotación óptica por moléculas, la paridad se conserva.

Para el átomo, sin embargo, no hay lateralidad geométrica alguna. Cualquier lateralidad en un gas de átomos tiene que ser impuesta desde fuera, quizás aplicando un campo magnético externo a los átomos o dándoles alguna orientación preferida a su espín. En ausencia de tal lateralidad impuesta externamente, los átomos se ven igual en un espejo a como se ven en la realidad. Supongamos que en tal medio se detecte una pequeña rotación óptica de un haz de luz polarizada linealmente. En un espejo se verían átomos idénticos a los átomos reales, dando lugar, no obstante, a una rotación óptica contraria a la observada en la realidad. Se violaría la conservación de la paridad. La rotación óptica en un gas atómico no está causada por la lateralidad en la geometría de los átomos, como sucede en un gas molecular, sino por la lateralidad incorporada en las leyes de la física que gobiernan la fuerza débil.

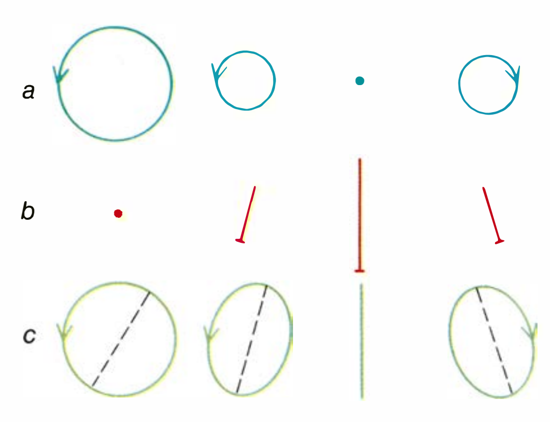
El ángulo de rotación óptica predicho por la teoría electrodébil es extremadamente pequeño, unos 10^{-5} grados en las circunstancias experimentales más favorables. El ángulo viene a ser igual al ángulo subtendido por la anchura de una aguja de coser vista a unos ocho kilómetros de distancia. No obstante, varios grupos experimentales han llevado a cabo exámenes sensibles que deberían detectar tales pequeñas rotaciones en un gas calentado de átomos de bismuto. Los primeros resultados fueron publicados, en 1976 y 1977, por investigadores de Oxford y Washington; eran unos resultados nulos: no se observó ningún efecto con la magnitud predicha. Durante cierto tiempo se especuló si no tendría que revisarse la teoría electrodébil de Weinberg y Salam.

Sin embargo, en 1978, L. M. Barkov y M. F. Zolotorev, también del Instituto de Física Nuclear de Novosibirsk, comunicaron la observación de rotaciones ópticas del tamaño y sentido apropiados en bismuto. Poco después de este resultado, investigadores del Centro del Acelerador Lineal de Stanford (SLAC) observaron un caso claro de violación de la paridad en la dispersión de electrones de alta energía por neutrones y protones. Aunque el descubrimiento estaba de acuerdo cuantitativamente con la teoría de Weinberg-Salam, no era estrictamente pertinente a la cuestión de conservación de la paridad en el átomo. Tras medidas repetidas con técnicas perfeccionadas, los investigadores de Oxford y Washington comunicaron finalmente rotaciones ópticas en bismuto de la magnitud esperada. El año pasado se detectó en Washington una rotación similar en plomo. Un cuarto grupo, en Moscú, que en 1980 había negado cualquier rotación en bismuto, acaba de observar una cuya magnitud concuerda con el resultado de Oxford.

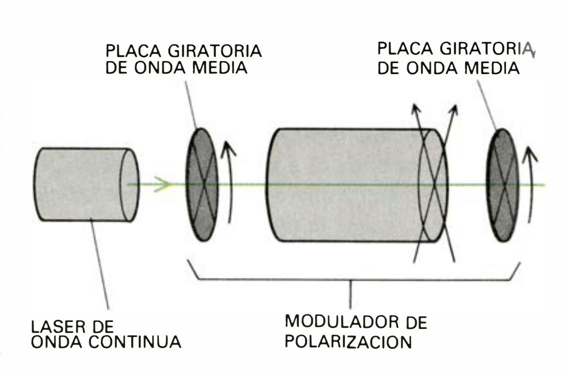
A pesar de estos hallazgos, en ensayos con rotación óptica se tiene que hacer frente a importantes efectos sistemáticos que tienden a simular los efectos de las violaciones reales de la paridad. De hecho, los resultados cuantitativos de los experimentos de Oxford y Moscú no están en completo acuerdo con los resultados de Novosibirsk, a pesar de que las tres experiencias miden las mismas propiedades del átomo de bismuto. (En Washington se hizo una medida diferente del átomo de bismuto; no podemos comparar, pues, directamente sus resultados.) Aunque no cabe la menor duda acerca de la existencia de rotación óptica en el bismuto, es aún algo prematura una conclusión teórica cuantitativa a partir de los ensayos.

En una segunda categoría, tres ensayos han intentado medir la violación de paridad en el átomo de hidrógeno. Por ser el número atómico Z del átomo de hidrógeno igual a 1, Z^3 valdrá también 1; no se producirá amplificación cúbica alguna de los efectos de violación de paridad. El efecto de la fuerza débil sobre el electrón en su estado energético más bajo en el átomo de hidrógeno es tan pequeño que el átomo debería ofrecer una conservación casi exacta de la paridad, en una parte en 10^{14}. Por otro lado, nos es dado explotar en el átomo de hidrógeno otro mecanismo que puede amplificar los efectos de violación de paridad. Dos estados energéticos excitados del átomo tienen casi la misma energía; trabajando con uno de los dos estados, que es metaestable, la violación de la paridad puede ser ampliada en varios órdenes de magnitud. Los ensayos con hidrógeno son, no obstante, mucho más difíciles que las correspondientes experiencias con átomos más pesados.

Hay una razón de peso para medir la violación de la paridad en el hidrógeno; la medición aportaría información sobre efectos de relativamente largo alcance de la fuerza débil, imposibles de determinar a partir de las experiencias con átomos más pesados. También proporcionaría mayor información sobre la estructura de la fuerza débil. Finalmente, se pueden realizar cálculos teóricos para el átomo de hidrógeno mucho más seguros y fidedignos de lo que se logra con átomos pesados. Los cálculos reflejan la simplicidad del átomo, en el que un solo electrón orbita alrededor del único protón del núcleo. Físicos experimentales de las universidades de Michigan, Washington y Yale han acometido las mediciones, pero hasta



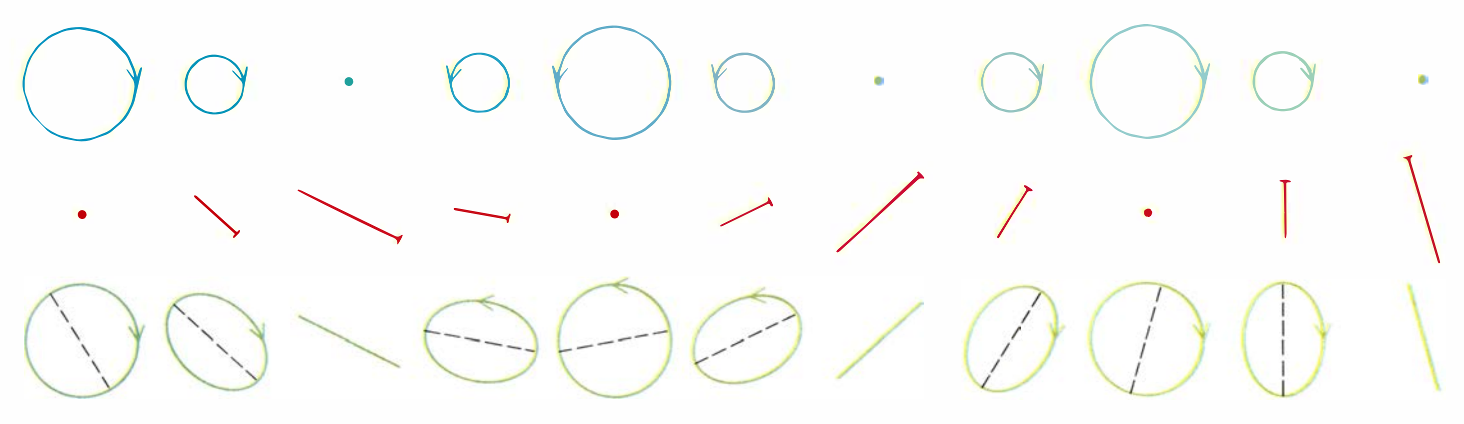
6. EL MODULADOR DE POLARIZACION cambia la luz monocromática, linealmente polarizada, emitida por un láser, en un haz cuya polari-



7. APARATO empleado por los autores del artículo y sus colegas para medir la violación de la paridad en el átomo de cesio. Consta de tres partes principales: la fuente de luz, la célula multipasos del haz que contiene la muestra y el montaje de detección. El láser y el modulador de polarización,

ahora no se dispone de técnicas suficientemente sensibles para someter a prueba la teoría electrodébil.

La tercera categoría de experiencias recoge dos ensayos realizados con cesio atómico por nosotros y nuestros colaboradores Jocelyne Guéna y Larry Hunter, en la École Normale Supérieure, otro ensayo con talio llevada a cabo por Eugene D. Commins y sus colegas, en la Universidad de California en Berkeley, y ensayos con cesio, todavía en marcha, en Michigan y en el Instituto Federal de Tecnología de Zurich. En todos estos ensayos, se excita una fracción de los átomos de un gas hasta un estadio energético más alto, mediante un haz de láser; se escoge una lateralidad preferida para la configuración experimental inicial. Los átomos despiden luz de fluorescencia en cuanto caen a un estado energético intermedio; se registra entonces las propiedades de polarización de tal luz de fluorescencia. Se cambia luego la configuración experimental a otra nueva, idéntica a la imagen especular de la primera configuración. Por ejemplo, se puede realizar el cambio invirtiendo un campo eléctrico o cambiando la lateralidad de la polari-



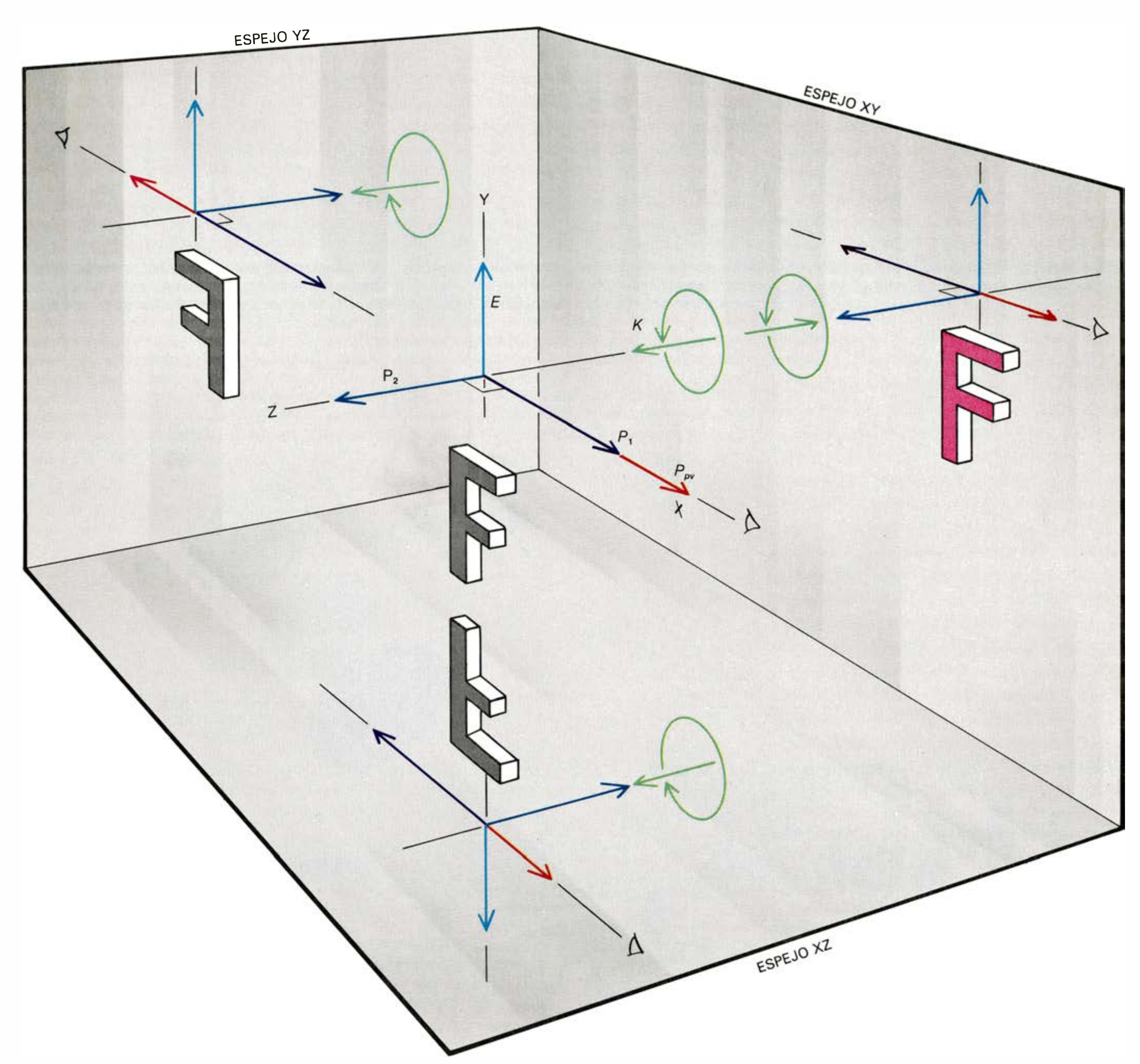
zación varía periódicamente con el tiempo. La polarización resultante (*c*) puede separarse en una componente circular (*a*) y una componente lineal (*b*); ambas oscilan con la misma frecuencia, si bien lo hacen desfasadas en un cuarto de período. De este modo, cuando una componente se anula, la otra está en su máximo. La componente lineal gira mientras oscila, pero la frecuencia de rotación difiere de la frecuencia de la oscilación de la amplitud.

que constituyen la fuente de luz, liberan luz verde monocromática, con una polarización que varía suavemente, hacia la célula multipasos del haz. En la célula, la luz rebota entre dos espejos y pasa alrededor de 100 veces a través del vapor de cesio, en una región limitada por las dos placas paralelas, cargadas, de un condensador. La luz excita una fracción pequeña de los átomos hasta un nivel de energía más alto; los átomos excitados tienen polarizado el espín. Cuando vuelven a un estado energético más bajo, emiten fluorescencia infrarroja, que es registrada y analizada por el montaje de detección. Los espejos se colocan dentro de la célula para eliminar la pérdida de intensidad que resultaría con cada paso del haz a través de una ventana de la célula; se protegen los espejos de los efectos corrosivos que produce el vapor de cesio mediante un amortiguador formado por gas de helio, enfriado por agua.

zación circular de la luz láser incidente. Se hace entonces una segunda medida de la fluorescencia atómica; si la segunda medida no se ajusta a la imagen especular de la primera, tenemos una prueba de la violación de la paridad en los átomos.

Podría parecer innecesariamente complicado trabajar con transiciones entre dos niveles energéticos, en vez de hacerlo con alguna propiedad de un nivel único. El efecto débil que podemos detectar aparece en el breve período durante el cual el átomo oscila entre los estados energéticos bajo y alto, antes de que permanezca durante un tiempo relativamente largo en el estado más alto. El átomo excitado lleva la señal de la interacción débil, adquirida durante las oscilaciones.

Cuando nos aprestemos a buscar efectos de la fuerza débil conviene reducir la magnitud efectiva de la fuerza electromagnética: esa fuerza puede imitar sistemáticamente el efecto de violación de paridad de la fuerza débil. Un modo práctico de minimizar el efecto de la fuerza electromagnética consiste en trabajar con una transición atómica "prohibida"; a una transición la llamamos prohibida cuando es muy improbable que ocurra en virtud de efectos electromagnéticos. En transiciones altamente prohibidas, la relación de la magnitud de la fuerza débil a la magnitud de la fuerza electromagnética aumenta substancialmente; sin embargo, incluso en tales transiciones, la fuerza electromagnética es la más fuerte.



8. ESPIN DE LOS ATOMOS EXCITADOS en el gas. Puede descomponerse en tres sumandos: P_1 (*morado*), P_2 (*azul oscuro*) y P_{pv} (*rojo*). P_1 y P_2 provienen en su totalidad de los efectos de la fuerza electromagnética; así, cuando la configuración experimental se reemplaza por su imagen especular, ambas componentes señalan en la misma dirección que sus imágenes especulares. P_{pv}, mucho menor en magnitud que las otras dos componentes, proviene de la interferencia entre la fuerza débil y la electromagnética. Esta componente viola la paridad del espín; cuando la configuración experimental se reemplaza por su imagen especular, P_{pv} se reemplaza por una flecha que señala en una dirección opuesta a la de la imagen especular de un vector axial. P_{pv} actúa como lo haría un verdadero vector si se conservase la paridad. El grado de polarización circular de la luz fluorescente emitida a lo largo del eje x depende de la magnitud total de la componente de espín allí; esta magnitud no es la misma para la configuración experimental inicial (*centro*) que para las configuraciones en los espejos. La lateralidad de la configuración inicial está establecida por la dirección del campo eléctrico E (*azul claro*), la dirección del haz de láser k (*verde*) y el sentido de la polarización circular del haz incidente. Una inversión adecuada de uno o dos de esos valores equivale a una reflexión especular en uno de los tres planos. La letra mayúscula F indica de qué manera cada reflexión transforma las condiciones experimentales iniciales.

En 1981, Commins y su grupo de Berkeley observaron violaciones de la paridad trabajando con una transición altamente prohibida en el átomo de talio. La incertidumbre en el resultado experimental era de un treinta por ciento. Mas, debido a la estructura algo compleja de talio, el valor teórico esperado de la violación de paridad podía calcularse también con una precisión probable de un 30 por ciento. Dentro de las incertidumbres experimentales y teóricas, los valores medido y calculado de la violación de la paridad están de acuerdo.

¿Por qué elegimos en nuestro ensayo el cesio? Sencillamente porque los cálculos necesarios de la estructura atómica son más fiables. Como todos los demás átomos de su columna en la tabla periódica, el cesio tiene un solo electrón en su capa más externa. Los restantes 54 electrones están bastante estrechamente ligados en una coraza electrónica que actúa de un modo parecido al que lo hace el átomo de un gas noble inerte. Por razones de cálculo, las perturbaciones de los electrones de la coraza pueden ser prácticamente despreciadas al considerar la órbita del electrón más externo. Además, el cesio encierra un número atómico moderadamente alto, y así la experiencia puede beneficiarse de la amplificación de los efectos de violación de la paridad asociada con el cubo del número atómico. La amplificación real es 10 veces menor en el cesio de lo que lo es en el bismuto, el plomo y el talio. El cesio es el más sencillo de los átomos pesados y el más pesado de los átomos sencillos.

En nuestro experimento, los átomos de cesio entran en la fase de gas evaporándose a partir de un depósito líquido en el interior de un tubo de vidrio cerrado. Se calientan a 200 grados Celsius, aproximadamente; se someten a un campo eléctrico entre las dos placas paralelas cargadas de un condensador. El campo eléctrico distorsiona cada átomo de cesio convirtiéndolo a una pequeña antena: la carga positiva se concentra a lo largo de una dirección en el átomo y la carga negativa se concentra a lo largo de la otra. La pequeña antena atómica puede resonar con las oscilaciones electromagnéticas de la luz del láser.

La oscilación del átomo de cesio, sin embargo, no la causan sólo efectos electromagnéticos. Además de la antena electromagnética inducida, el átomo posee una segunda antena, mucho menos sensible, asociada a la fuerza débil. Cuando las dos antenas son excitadas por las vibraciones resonantes de la luz láser incidente, pueden absorber ambas potencia del láser. La potencia absorbida no equivale a la suma de las potencias que serían absorbidas por las dos antenas si se vieran excitadas por separado. Antes bien, la potencia total absorbida depende también de la magnitud de la interferencia entre ambas antenas. La señal de violación de la paridad que observamos en nuestra experiencia depende de la magnitud de la interferencia.

Cada electrón tiene un espín similar al espín de un giróscopo. En física, el espín se representa por una flecha llamada vector axial, dirigido a lo largo del eje del objeto que gira, en otras palabras, perpendicular al plano del movimiento de rotación. La dirección del vector axial está fijada por una convención, la de la llamada regla de la mano derecha: el vector axial apunta en la misma dirección en que el pulgar derecho señalaría si los dedos de la mano derecha se enroscasen en la dirección del espín. Notemos que, de acuerdo con la regla de la mano derecha, el espín del electrón emitido en la desintegración beta señala en el sentido opuesto al de su movimiento hacia adelante.

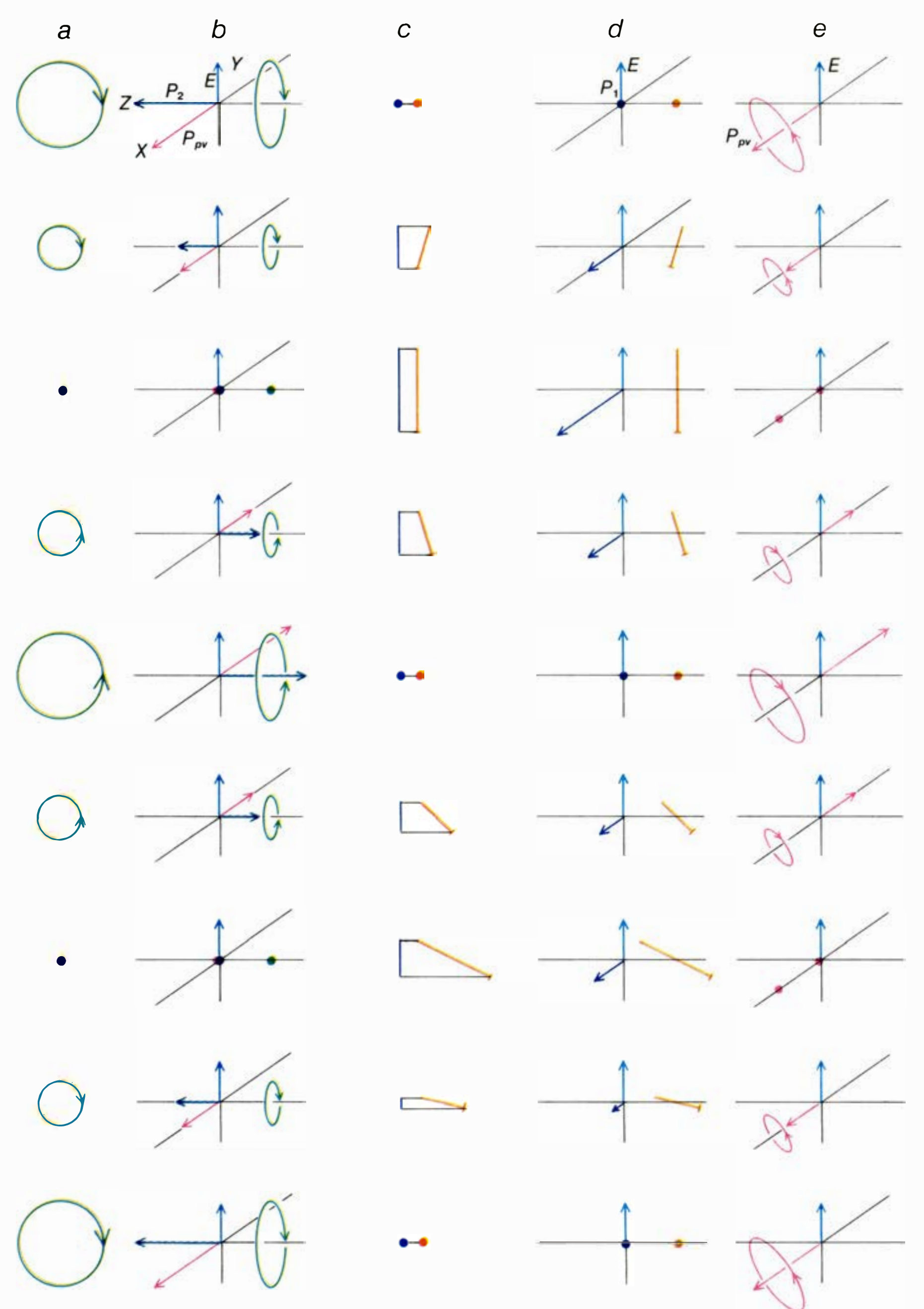
Un rasgo característico de la fuerza débil es que distorsiona la órbita electrónica de modo tal que tiende a alinear el espín del electrón con su velocidad. Sucede entonces que la antena atómica asociada con la fuerza débil está dirigida a lo largo del espín electrónico. En un gas de átomos, los vectores axiales que representan los espines electrónicos señalan al azar en todas las direcciones; el espín promedio es cero. La misma propiedad se aplica, inicialmente, a los electrones en los átomos de cesio sometidos al campo eléctrico de nuestra experiencia. La dirección de la antena atómica asociada con la fuerza electromagnética depende, sin embargo, de la dirección del campo eléctrico impuesto por el condensador. La interferencia entre los dos tipos de antena depende del ángulo entre ellos; ese ángulo depende, a su vez, de la dirección del campo eléctrico y de la dirección y polarización de la luz láser que fuerza las oscilaciones de las antenas y de la dirección de los espines electrónicos. Por ser aleatorios los espines, pero no las otras propiedades, los átomos cuyos espines electrónicos están orientados en ciertas direcciones se excitan preferentemente por el haz de láser hasta un estado de energía más alto. Como resultado, los átomos que llegan a excitarse tienen polarizado el espín. En otras palabras, el espín electrónico promedio de los átomos excitados no es nulo: apunta en cierta dirección.

El vector axial que representa el espín electrónico total de los átomos en el estado excitado puede descomponerse en tres sumandos [*véase la figura 8*]. Dos de las componentes, designadas por P_1 y P_2, son perpendiculares entre sí y también a la dirección del campo eléctrico; ambas se originan sólo de la fuerza electromagnética. La tercera componente proviene de la interferencia entre la fuerza electromagnética y la fuerza débil. Se la designa P_{pv}, por ser la componente del espín que viola la paridad. P_{pv} apunta a lo largo del mismo eje que la componente P_1.

Una vez excitado el átomo, su transición de vuelta al estado de energía inicial sigue prohibida, y, en cambio, permitida la transición a un estado de energía intermedia. En tal proceso, el átomo emite un fotón de radiación infrarroja cuya energía es igual a la diferencia en energía entre el estado excitado y el estado intermedio. Como el momento angular del sistema tiene que conservarse, el fotón emitido se lleva parte del momento angular del átomo con espín. En otras palabras, parte de la polarización del espín de los átomos excitados se convierte en polarización de la luz de fluorescencia; así puede inferirse la polarización del espín. De modo más preciso, el grado de polarización circular de la luz de fluorescencia emitida en una dirección dada es proporcional al tamaño de la componente de espín atómico en esa dirección. Medimos el grado de polarización circular en la luz de fluorescencia emitida a lo largo del eje definido por las dos componentes de espín P_1 y P_{pv}. De aquí, obtenemos la suma de las dos componentes, P_1 y P_{pv}.

Nos interesa sobre todo la magnitud de P_{pv}. ¿Cómo distinguirla de P_1, que está dirigida a lo largo del mismo eje? Y por ende, ¿cuál es la diferencia entre las componentes P_1, P_2 y P_{pv}? Para contestar a esas preguntas conviene recordar qué sucede cuando un vector axial se refleja en un espejo.

Cuando se observa un disco en rotación en un espejo paralelo al eje de espín del disco, el disco y su imagen especular aparecen girando en direcciones opuestas. Si el espejo es perpendicular al eje del disco, el disco y su imagen se nos ofrecerán girando en la misma dirección. Como el vector axial se dirige a lo largo del eje del espín, la reflexión en un espejo paralelo al vector axial invierte su dirección, mientras la reflexión en un espejo perpendicular al vector axial deja su dirección inalterada. Las transformaciones son exactamente opuestas a las transformaciones



9. MAGNITUD DE LAS COMPONENTES DE ESPIN P_2 (*azul oscuro*) y P_{pv} (*rojo*) (*b*). Dicha magnitud es proporcional a la parte del haz láser incidente que está polarizado circularmente (*verde*) (*a*). La magnitud de la componente de espín P_1 (*morado*) (*d*) es proporcional a la proyección, a lo largo del eje del campo eléctrico de la polarización lineal del haz láser (*naranja*) (*c*). La señal se obtiene a partir de la polarización circular de la luz de fluorescencia (*rojo*), observada a lo largo del eje *x* [*véase el esquema inicial en b*]. Esa parte de la polarización circular es proporcional a la polarización circular del haz láser (*e*).

que sufre un verdadero vector en reflexiones especulares [*véanse las figuras 1, 2* y *3*].

Podemos ya establecer la diferencia entre P_1 y P_{pv}. Cuando sustituimos la configuración experimental que da lugar a P_1 por una segunda configuración, que coincide con la imagen especular de la primera, la dirección de la contrapartida de P_1 en el espejo corresponde a la imagen especular de un vector axial. En otras palabras, el vector axial al que da lugar la configuración original se refleja tal como debería ocurrir cuando se conserva la paridad. Por otra parte, la dirección de la contrapartida de P_{pv}, en la segunda configuración, es exactamente opuesta a la de la imagen especular de un vector axial. Al generarse P_{pv} a través de una fuerza que no conserva la paridad, actúa como lo haría un verdadero vector si en el ensayo se conservase la paridad.

En la práctica hay otros medios más para distinguir las componentes de espín P_1, P_2 y P_{pv}. Su magnitud y su dirección están gobernados por la polarización del haz de láser que proporciona energía; vale, pues, la pena examinar cómo se establece la polarización del haz. Antes de que éste alcance el vapor de cesio, atraviesa el llamado modulador de polarización. La polarización de la luz que emerge de ese dispositivo oscila con el tiempo desde polarización circular dextrógira total hasta polarización lineal completa, polarización circular levógira completa y vuelta, de nuevo, a la inicial. Además, la dirección de la polarización lineal gira suavemente con el tiempo, aunque con un período diferente del invertido en la oscilación entre polarización circular levógira y dextrógira [*véase la figura 6*]. Importa destacar que la polarización circular del haz de láser incidente confiere lateralidad a la configuración experimental. Recordemos que, al moverse en un círculo el vector campo eléctrico del láser, el haz de luz se mueve hacia adelante, por lo que la punta del vector traza una hélice levógira o dextrógira. La oscilación de la luz entre la polarización circular levógira y dextrógira es uno de los modos que permiten invertir periódicamente la lateralidad de la configuración experimental.

Esas modulaciones del haz de láser modifican las componentes de espín del gas atómico en términos específicos que sirven para designar las componentes. Las componentes P_2 y P_{pv} son, ambas, proporcionales a la cantidad de polarización circular del haz incidente; se modulan por el cambio periódico de esa polarización. La componente P_1 es proporcional a la proyección de la polarización lineal del haz a lo largo del eje paralelo a la dirección del campo eléctrico. Debido a la rotación de la polarización lineal, P_1 se modula con un ritmo que no se ajusta a las oscilaciones de P_2 y P_{pv} [*véase la figura 9*].

De tres sencillas maneras reemplazamos nuestra configuración experimental por su imagen especular; corresponden a tres planos mutuamente perpendiculares de reflexión especular espacial. Primero, tal como hemos examinado, el modulador de polarización puede invertir la polarización circular del haz incidente, mientras deja inalteradas las direcciones del haz y el campo eléctrico. La inversión corresponde a una reflexión en un espejo perpendicular a P_1 y P_{pv}. Segundo, tanto la polarización circular como el campo eléctrico establecido por el condensador pueden invertirse, lo que corresponde a una reflexión en un espejo perpendicular al campo eléctrico. Tercero, la dirección del haz láser puede invertirse, mientras se mantienen el campo eléctrico del condensador y la dirección de rotación absoluta del vector que representa el campo eléctrico del láser. El cambio de

dirección modifica la lateralidad del haz láser y, por tanto, corresponde a una reflexión perpendicular al haz.

En la práctica, invertimos continuamente la dirección del haz láser con dos espejos esféricos. Cada espejo está montado sobre un lado de la vasija de vidrio que encierra el vapor de cesio. El haz láser va y viene unas 50 a 60 veces entre los espejos antes de que escape a través de la abertura de entrada, en uno de ellos. El dispositivo tiene una ventaja: la señal que viola la paridad queda amplificada casi en proporción al número de pasos del haz. La componente no deseada de espín, P_1, por otra parte, cambia su dirección con cada paso; por cuya razón los pasos múltiples tienden a cancelarla. A pesar de ello no puede ser eliminada del todo; se mantiene unas 30 veces mayor que la componente P_{pv}.

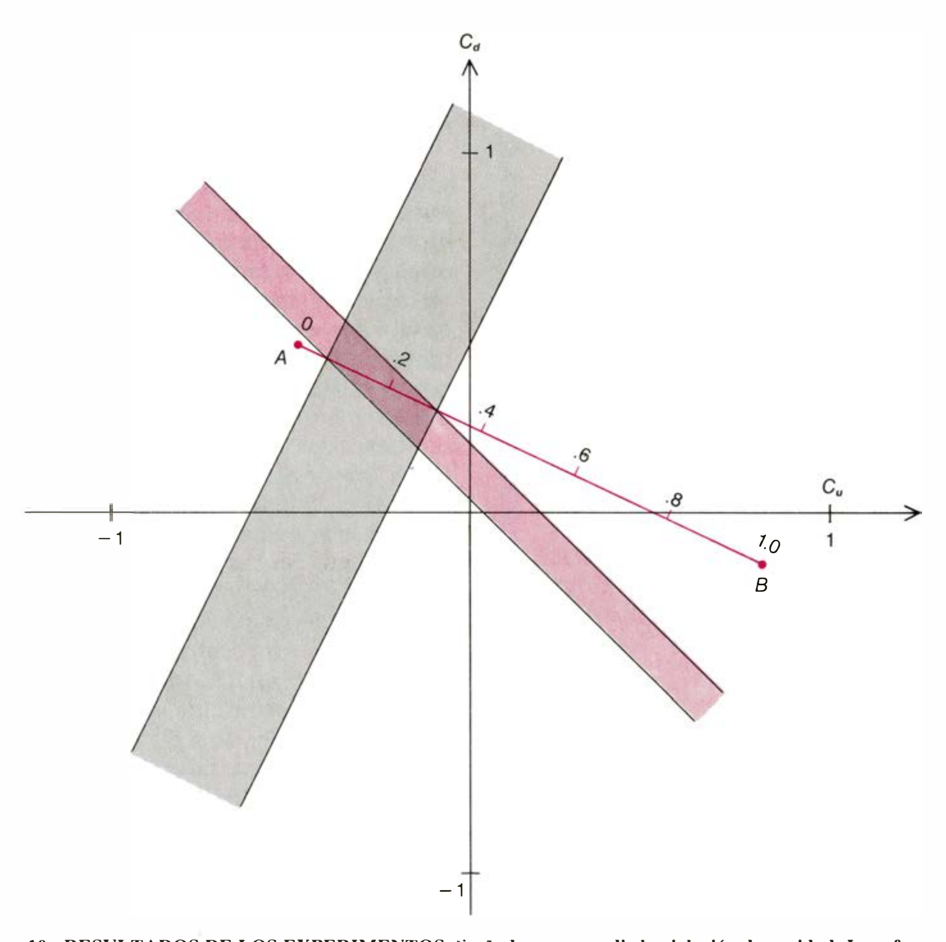
Pusimos todo el empeño en el análisis, reducción y control de los efectos sistemáticos. En general, las imperfecciones de la geometría o las inversiones de la lateralidad del experimento engendran errores sistemáticos. Por ejemplo, al invertir la polarización circular del haz de láser con el modulador, las otras propiedades del haz, tal como su intensidad y su grado de polarización lineal, tienen que permanecer inalteradas. Basta un cambio espurio de sólo un uno por ciento para simular la esperada violación de la paridad. Lo inconveniente del caso estriba en que las imperfecciones de este tipo en los moduladores de polarización existentes suelen ser del orden de varias unidades por ciento. Para superar el problema, tuvimos que diseñar un nuevo tipo de modulador, capaz de garantizar una inversión pura de la polarización dentro de los límites de una parte en 10^4.

Otra dificultad sistemática reside en las imperfecciones de los espejos, que pueden cambiar ligeramente la polarización de la luz cuando ésta se refleja; efecto que simula también la violación de paridad. No conocíamos ningún industrial que fabricase espejos con la tolerancia que necesitábamos; tuvimos que escoger entre una veintena disponible, de la que dos resultaron adecuados. Una vez que se hubieron instalado, girábamos el espejo de retorno de vez en cuando, en un ángulo de 90 grados en su plano. El ajuste periódico invertía cualquier efecto sistemático y, en promedio, lo cancelaba.

Había muchas otras imperfecciones en potencia a considerar: campos eléctricos incidentales o inversión imprecisa del haz en la vasija de vidrio. Medimos esas imperfecciones usando los propios átomos como prueba. Las incluimos todas en un modelo teórico que daba una estimación cuantitativa de los posibles efectos sistemáticos. Tal estimación desempeñaba un papel crucial para determinar la confianza que se podía depositar en la medida. En nuestro afán por comprender y controlar efectos sistemáticos gastamos más de tres años. Pretendíamos mantener todos los efectos sistemáticos conocidos por debajo de unas pocas unidades por ciento del valor medido. Logramos nuestro propósito: no nos vimos en la obligación de ninguna corrección significativa al valor medido bruto de la violación de la paridad.

Frente a posibles efectos sistemáticos inesperados, no recogidos en el modelo, no quedaba otra alternativa realista que verificar que los datos proporcionados por la experiencia satisficieran cualesquiera condiciones que demandase el sentido común. Sucede, por ejemplo, en nuestro experimento que la señal de violación de paridad se obtiene simultáneamente a través de dos canales independientes, que corresponden a dos frecuencias de modulación. Hemos comprobado que las señales transportadas por los dos canales están de acuerdo. Mientras se medía la paridad, se llevaron a cabo también dos mediciones nulas: los dos resultados fueron nulos tal como se esperaba. Se realizaron exámenes estadísticos para verificar la ausencia de algunas correlaciones no deseadas. Finalmente, a mitad de nuestra adquisición de datos, nos dimos cuenta, de pronto, que las dos placas giratorias del modulador de polarización podían –¿quién sabe?– introducir alguna preferencia espuria entre izquierda y derecha. Sus rotaciones fueron invertidas para todo el conjunto de datos subsiguiente; los resultados anteriores y posteriores a la inversión concordaron.

Para calibrar nuestras mediciones de la polarización electrónica, medíamos



10. RESULTADOS DE LOS EXPERIMENTOS diseñados para medir la violación de paridad. Los ofrecemos aquí sobre los mismos ejes. El eje C_u representa la carga débil del quark arriba (*u*); el eje C_d representa la carga débil del quark abajo (*d*). Los experimentos de alta energía realizados en el Centro del Acelerador Lineal de Stanford (SLAC) confinan los valores posibles de las dos cargas a una franja lineal (*gris*). Los ensayos atómicos de baja energía desarrolladas por los autores y sus colegas confinan los valores posibles a otra franja lineal, casi perpendicular a la primera (*color*). Tomados juntos, ambos experimentos ponen límites estrechos a los valores de C_u y C_d. Los valores predichos por la teoría electrodébil estándar corresponden a un punto sobre el segmento de línea en color. Ese punto, sin embargo, está indeterminado en la teoría, pues depende del valor asignado a cierta variable libre, o arbitraria, en la teoría. La teoría estándar es coherente con los puntos sobre la intersección de las franjas para valores de esa variable cercanos a 0,2. Los resultados de los ensayos que determinan la variable convergen en 0,23.

la gran componente P_2 con el mismo aparato que medía P_{pv}. Se aplicaba un campo magnético durante un corto tiempo, a lo largo del eje paralelo al campo eléctrico. El campo magnético giraba P_2 de modo tal que podía detectarse a partir de la misma dirección que P_{pv}. Lo que realmente medimos en nuestra experiencia era la relación de P_{pv} a P_2, cifrada en torno a 10^{-5}. La relación daba directamente la cantidad observada de asimetría izquierda-derecha; era el análogo del pequeño ángulo de rotación óptica observado en los experimentos con bismuto. A partir de estudios espectroscópicos se conocía el valor de P_2; cabía, pues, calcular el valor de P_{pv} a partir de la relación medida de P_{pv} a P_2.

Dadas las incertidumbres de la experiencia y del cálculo teórico, el primer resultado de nuestra experiencia está en excelente acuerdo con la teoría de Weinberg-Salam. La incertidumbre estadística es del 16 por ciento mientras la incertidumbre sistemática es del 8 por ciento. Claude Bouchiat y sus colegas han calculado la violación de la paridad predicha por la teoría: la incertidumbre en su predicción no llega al 15 por ciento.

El año pasado comprobamos el resultado con una versión modificada de nuestro experimento en el cesio. Una y otra versión del ensayo difieren en las orientaciones relativas de los espines electrónico y nuclear de los átomos excitados por el haz láser. Esta sutil diferencia práctica da lugar, en principio, a una gran diferencia. En las dos versiones, la violación de la paridad se detecta a través de cantidades físicas diferentes; la mayor parte de las estimaciones de errores sistemáticos potenciales se basan a su vez en mediciones distintas. No obstante, ambos resultados experimentales tienen aproximadamente igual precisión; también están de acuerdo entre sí y con el valor teóricamente predicho, que queda entre ellos. Al combinar los dos resultados, la incertidumbre estadística se reduce del 16 al 11 por ciento; la incertidumbre sistemática permanece alrededor del 8 por ciento.

A pesar del acuerdo de nuestro ensayo con la formulación actual de la teoría estándar de la fuerza electrodébil, resulta instructivo interpretar nuestro resultado con independencia de dicha teoría. En átomos pesados, la componente de la fuerza débil a destacar entre el electrón y el núcleo es proporcional a la carga débil nuclear, que es suma de todas las cargas débiles que forman el núcleo. El núcleo consta de protones y neutrones; éstos, a su vez, están formados por quarks *u* (arriba) y *d* (abajo). Ensayos como el nuestro determinan el valor de la carga débil nuclear, pero ese valor no basta para determinar las cargas débiles de los quarks *u* y *d* por separado. Antes bien, sólo impone una relación entre los valores posibles de las cargas débiles de los quarks. Si las cargas de los quarks se representan sobre ejes independientes, en un sistema de coordenadas rectangulares, el punto que corresponde a su valor correcto tiene que estar a lo largo de una línea recta. Solamente la línea, y no la posición del punto a lo largo de ella, queda determinada por una medida experimental de la carga nuclear débil.

Los resultados del experimento de altas energías desarrollado en el SLAC pueden representarse sobre el mismo conjunto de ejes coordenados; pero conducen a una relación lineal bastante diferente. Si los resultados del experimento de altas energías se combinan con los resultados del trabajo en física atómica, los valores de las cargas débiles de los quarks quedan confinados a la intersección de las dos líneas. En principio, por tanto, los valores están determinados. En la práctica, debido a la incertidumbre, tanto experimental como en la interpretación teórica, cada línea es una franja de anchura finita. No obstante, la intersección de las dos franjas representa aún sólo un conjunto pequeño de valores posibles. Resulta que esos valores restringidos, aunque se deducen, virtualmente, sin hipótesis teóricas, son coherentes también con la teoría de Weinberg y Salam.

El apoyo mutuo de dos tipos de ensayos del todo distintos para lograr una conclusión definida, inalcanzable por cualquiera de ellos aisladamente, ilustra la naturaleza cooperativa y complementaria de esta empresa. Cada uno de los difíciles experimentos que hemos abordado constituye un reto a muchos problemas técnicos, una empresa cuya evolución nadie puede adivinar, ni domeñar mientras avanza. El hecho de que se exploren diferentes caminos y que grupos distintos compitan tiene el sabor del estímulo recíproco, al par que es garantía de seguridad frente a resultados espurios. Considerando el interés de esos estudios, a pesar de sus dificultades, quizá los experimentos en marcha consigan una mejor precisión y continúen proporcionando nueva información sobre la naturaleza de la fuerza débil.



Ecología del bosque esclerófilo mediterráneo

Los estudios sobre la estructura y el funcionamiento del encinar sugieren nuevos aspectos de los mecanismos adaptativos con los que la vegetación responde a las condiciones específicas de la vida en la región mediterránea

Antoni Escarré, Carlos Gracia, Ferran Rodà y Jaume Terradas

El bosque mediterráneo ha sido mucho menos estudiado que los bosques caducifolios de invierno o los bosques boreales de coníferas, que son sus vecinos hacia el norte y que le superan ampliamente en extensión y valor económico. Sin embargo, el bosque esclerófilo tiene gran importancia local en la Península Ibérica y otros territorios de la región mediterránea. En los últimos años, los ecosistemas mediterráneos han sido objeto de numerosos estudios ecológicos, pero éstos se han centrado principalmente en los ecosistemas dominados por vegetación arbustiva esclerófila (de hoja dura y coriácea), como el chaparral de California, el *mallee* australiano, el *fynbos* sudafricano, el matorral chileno, la *frygana* griega, o las maquias y garrigas del sur de Europa. Son todavía escasos los trabajos sobre los bosques esclerófilos mediterráneos, a pesar de que manifiestan una personalidad acusada, con características peculiares muy notables que se han desarrollado en respuesta al doble estrés climático, sequía estival y frío invernal, y en respuesta también, quizás, a la escasez de nutrientes. En la Península Ibérica, junto a la importancia histórica de la explotación de la biomasa del bosque esclerófilo con fines energéticos, hoy de nuevo en auge, conviene destacar su valor estético, paisajístico y protector del suelo, y su carácter de ecosistema más complejo y maduro posible en la mayor parte del territorio.

En la región mediterránea encontramos dos grandes tipos de bosques esclerófilos, ambos de hoja perenne: los dominados por coníferas, cuyas hojas son aciculares o escuamiformes, y los constituidos por especies planifolias o de hoja ancha. En el presente artículo nos ocuparemos sólo de estos últimos, que en la Península Ibérica están representados principalmente por los bosques de encina (*Quercus ilex*) y de alcornoque (*Quercus suber*). Ambas especies han sido investigadas con cierta intensidad desde el punto de vista ecofisiológico y se conoce bastante bien sus pautas de comportamiento por lo que se refiere al intercambio de gases (fotosíntesis y transpiración). Se ha estudiado también otras especies arbóreas esclerófilas planifolias como el olivo (*Olea europaea*) y el algarrobo (*Ceratonia siliqua*), así como algunos de los arbustos y pequeños árboles más destacados de entre los que acompañan a los citados, por ejemplo el madroño (*Arbutus unedo*). Hoy sabemos ya que todos esos vegetales están más o menos adaptados al doble estrés mencionado antes: la duración del período de temperaturas subóptimas en invierno y la del período de suministro insuficiente de agua para compensar la evapotranspiración con los estomas (poros microscópicos de la hoja) abiertos, en verano. Ambos factores son determinantes para establecer el balance entre producción y respiración.

Los episodios de temperaturas especialmente bajas que se dan de vez en cuando en la región parecen desempeñar, además, un papel decisivo en la distribución de estos vegetales. Mientras la encina se muestra bastante resistente y no presenta lesiones hasta mínimas entre −15 y −25 grados Celsius, el olivo, el alcornoque, el lentisco (*Pistacia lentiscus*) o la coscoja (*Quercus coccifera*) se resienten en las hojas entre −12 y −14 grados C, y, entre −6 y −8 grados C, el algarrobo, el mirto (*Myrtus communis*) o las adelfas (*Nerium oleander*). La actividad metabólica se reduce de modo significativo por debajo de los 10 grados C, aunque la absorción de dióxido de carbono prosigue, si bien disminuida, por debajo de los 0 grados C. También la capacidad fotosintética (a temperaturas óptimas) es bastante menor en invierno. En ambos aspectos, la encina goza de una notable resistencia, si la comparamos con el olivo y otras especies.

En verano, la absorción de CO_2 viene limitada por la necesidad de reducir las pérdidas de agua cerrando los estomas. En la encina, madroño, coscoja, labiérnago (*Phyllirea media*) y otras especies similares, durante el período de estrés atmosférico, a mediodía, se produce un cierre parcial o casi completo de los estomas, que pueden abrirse otra vez por la tarde. En la encina, la coscoja y el labiérnago, O. C. Lange y colaboradores han medido conductancias de vapor de agua, en las hojas, a mediodía, de sólo el 15 al 20

1. CUENCA EXPERIMENTAL de 4,3 hectáreas que los autores han equipado en el macizo del Montseny para el estudio de los balances de agua y nutrientes en un encinar no perturbado durante las últimas décadas. Las pequeñas cuencas experimentales son una poderosa herramienta para investigar el funcionamiento de los ecosistemas terrestres. Debidamente seleccionadas, instrumentadas y mantenidas, tales cuencas permiten establecer el balance entre los nutrientes que entran en la cuenca con el agua de lluvia y los que salen de ella con el agua de avenamiento. Sumado a estudios paralelos sobre el uso de la luz, la economía del agua y la circulación de los nutrientes en el seno de los ecosistemas de la cuenca, este enfoque permite analizar en detalle el comportamiento de los sistemas terrestres en condiciones naturales y tras perturbaciones experimentales. En la fotografía se aprecia la estación de aforo, con un vertedero en V y un limnígrafo que registra de modo continuo el caudal de agua que sale de la cuenca. Otros instrumentos visibles son el pluviómetro, el pluviógrafo, los colectores de lluvia para análisis y un colector automático de agua del arroyo. Estudios similares se están realizando también en otra cuenca de encinar denominada Barranc de l'Avic, en las montañas de Prades a 120 kilómetros al suroeste del Montseny. Se espera obtener con todo ello un mejor conocimiento de la ecología del bosque esclerófilo mediterráneo.

por ciento de las máximas que se obtienen por la mañana; por la tarde, ese valor ascendía hasta el 30 o el 40 por ciento del máximo.

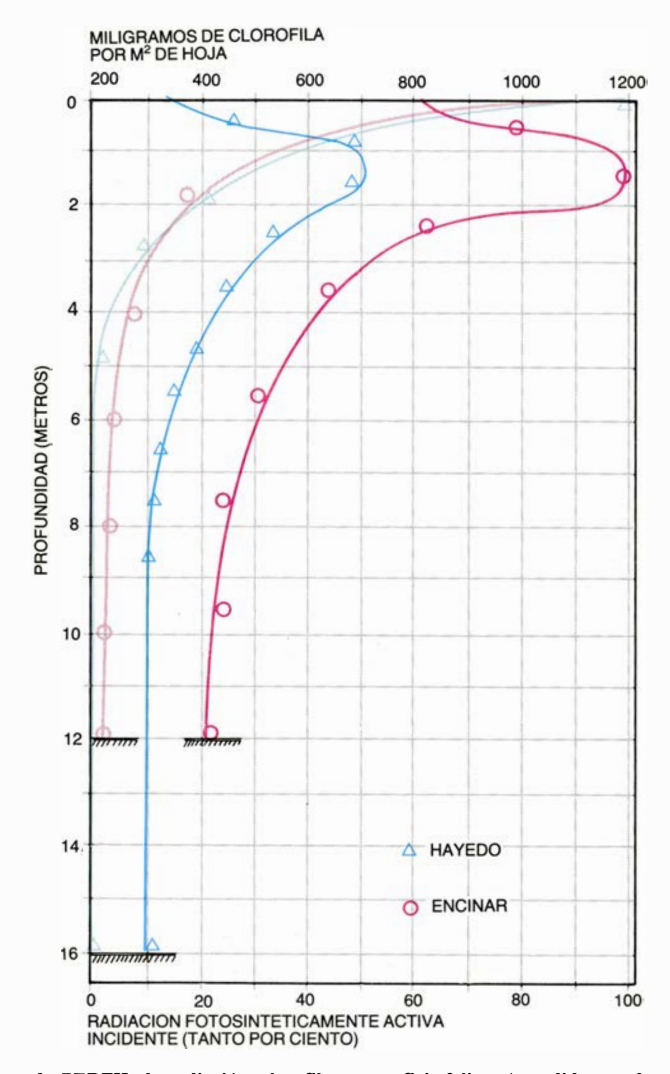
También en las relaciones hídricas observamos importantes diferencias de estrategia entre las diversas especies: por ejemplo, en cuanto a la existencia de la depresión de mediodía –ni la adelfa ni la zarzaparrilla (*Smilax aspera*) la presentan habitualmente– y a la sensibilidad en la respuesta de los estomas. En este último aspecto, se ha puesto de manifiesto que, además del control interno debido a las variaciones del contenido hídrico, y del control externo impuesto por las condiciones de luz, temperatura y humedad, puede haber en la encina componentes endógenos, posiblemente hormonas, que explicarían la reapertura de los estomas por la tarde, cuando todavía el estrés es máximo.

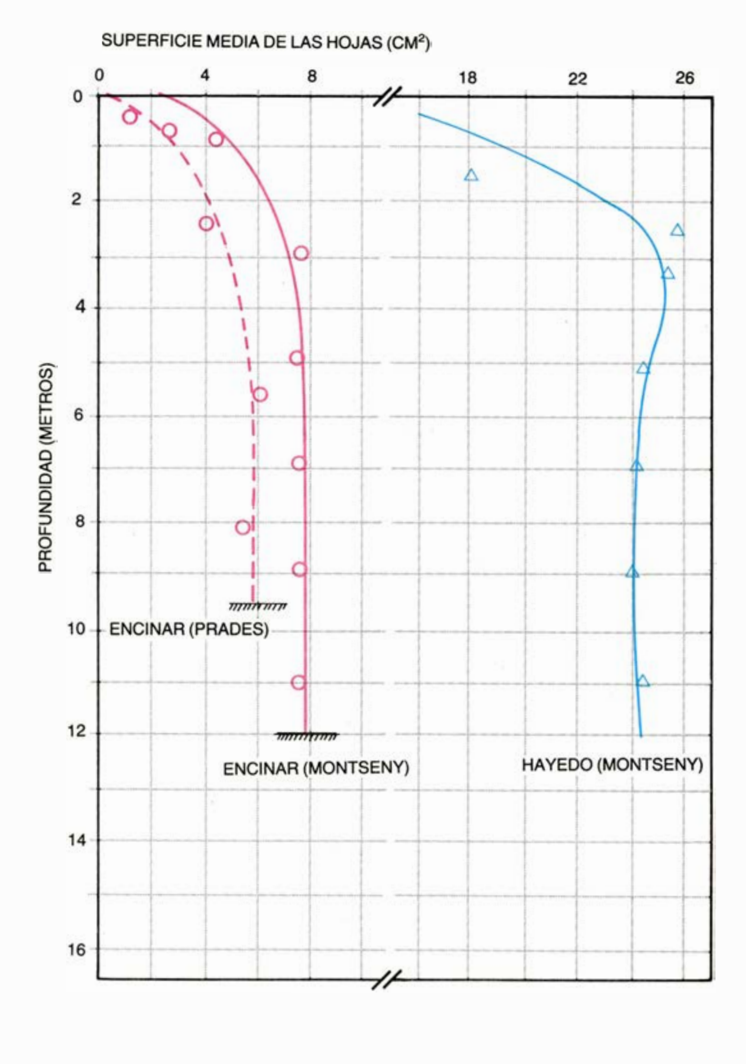
Pero un bosque es una estructura sumamente compleja, y estas consideraciones generales sobre las adaptaciones a nivel de especie reflejan pobremente la diversidad de matices con que esta estructura responde a las cambiantes condiciones del medio. Desgraciadamente, los bosques esclerófilos mediterráneos han sido poco estudiados a nivel del ecosistema. Antes de nuestros propios trabajos en los encinares de Prades y del Montseny, en la Cordillera Prelitoral Catalana, sólo el encinar de Le Rouquet en los alrededores de Montpellier había sido investigado en detalle por P. Lossaint, M. Rapp y colaboradores.

Para mejor comprender el funcionamiento del ecosistema fijémonos, en primer lugar, en su organización vertical. Un bosque puede interpretarse como un sistema basado en un importante transporte vertical y cuya estructura responde a las necesidades de este transporte y a todo un conjunto de gradientes ambientales: luz, humedad, temperatura y otros. En el bosque esclerófilo mediterráneo, algunas de estas pautas de organización vertical presentan peculiaridades interesantes.

En cualquier bosque cerrado, las hojas de los niveles superiores quedan mucho más expuestas a la radiación incidente que las de los niveles inferiores, a las que llega una luz tamizada por las que están situadas más arriba. Se establece así un gradiente de radiación. La luz, al atravesar el bosque, se va extinguiendo según una razón exponencial negativa, llamada ley de Lambert-Beer. Pero lo que ahora nos interesa es saber que, asociado al gradiente de luz resultante, se desarrolla un gradiente de formas y tamaños foliares.

Las hojas superiores, más expuestas a la radiación, se calientan intensamente sobre todo en los meses de verano,





2. PERFIL de radiación, clorofila y superficie foliar. A medida que descendemos en el vuelo del bosque, profundizando desde el nivel superior de las copas, la luz se va extinguiendo (*tonos claros*). Se pone de manifiesto la estrecha relación entre luz y clorofila presente (*tonos oscuros*). Por encima de los dos metros, cuando la radiación supera el 30 por ciento de la incidente, esta relación deja de ser válida porque el exceso de luz satura los fotosistemas y la clorofila se degrada con mayor rapidez tanto en el encinar (*rojo*) como en el hayedo (*azul*). La superficie de las hojas del nivel superior del bosque es menor que las de los niveles inferiores. Esta pauta no es exclusiva del bosque mediterráneo; se relaciona con el ambiente más seco y más expuesto a la radiación a que se ven sometidas las hojas del nivel superior. En el clima menos lluvioso de Prades, las hojas son menores que en el Montseny.

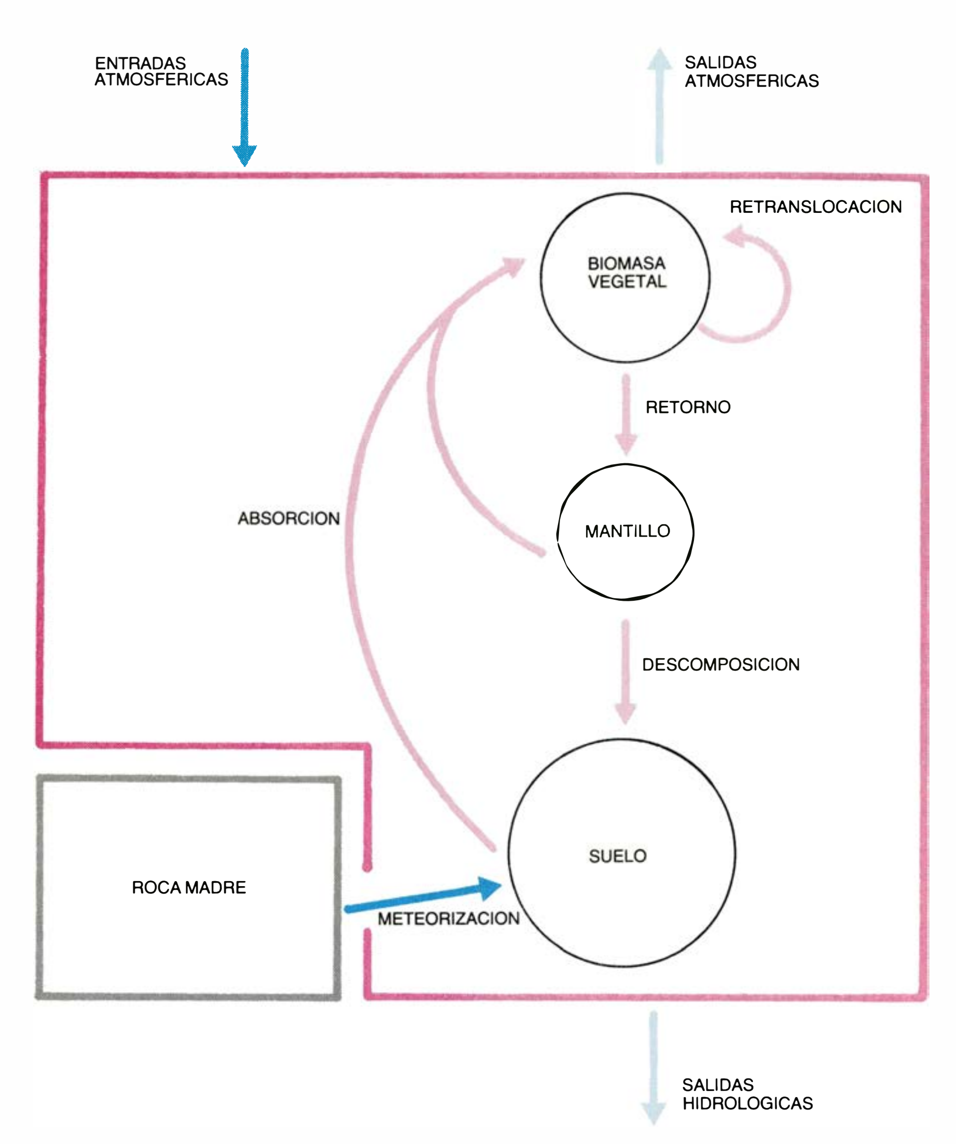
cuando la sequía estival, que coincide con la máxima intensidad de radiación incidente, obliga a la planta a reducir su transpiración. Estas hojas alcanzan temperaturas considerablemente superiores a la del aire. La disminución de la superficie de cada hoja y la lobulación del borde son dos modificaciones que facilitan el intercambio de calor con el ambiente. Disminuyendo la superficie y recortándola, mengua el espesor de la capa límite del aire en contacto con la hoja y se favorece la difusión del calor. Las hojas superiores, más pequeñas y generalmente más recortadas, representan una adaptación a las condiciones que imperan en la parte alta de la bóveda forestal. En el encinar, a medida que descendemos por las copas, las hojas doblan su tamaño y adoptan formas más enteras, con el borde menos lobulado.

En las plantas, la clorofila es la molécula especializada en la captación de la radiación electromagnética. Asociado al gradiente de radiación en la bóveda del bosque, existe un gradiente en la repartición de este pigmento. La concentración de clorofila es máxima justo por debajo de la primera capa de hojas, a un metro o dos por debajo del nivel superior del bosque. Por encima de este nivel, la luz satura los fotosistemas. (Aproximadamente el 30 por ciento de la radiación que incide en un día claro resulta saturante.) Este hecho y la elevada temperatura foliar contribuyen a degradar rápidamente la lábil molécula de clorofila en la parte superior de la copa de los árboles. Por debajo de los dos primeros metros, la luz se ha extinguido hasta un punto tal que ya no resulta saturante y aparece un máximo en la concentración de clorofila foliar. A partir de este punto, la concentración en las hojas decrece progresivamente hasta llegar al suelo, en consonancia con el proceso que sigue la luz, poniéndose de manifiesto la estrecha relación entre luz y clorofila.

El máximo subsuperficial de clorofila corresponde en el encinar a una concentración de unos 900 miligramos por cada metro cuadrado (en adelante, mg m^{-2}) de superficie foliar. Puesto que unos 400 mg m^{-2} bastan para absorber más del 99 por ciento de la radiación que llega a la hoja, resulta sorprendente encontrar estos valores tan elevados en la encina y en otras muchas plantas del bosque mediterráneo. Lentiscos, labiérnagos y otras especies presentan igualmente concentraciones elevadas de clorofila. La clave de la explicación hay que buscarla en la esclerofilia que caracteriza al bosque mediterráneo. Las hojas, que se proveen de gruesas cutículas e importantes parénquimas para resistir la deshidratación, se tornan prácticamente opacas a la luz. La luz que llega por un lado de la hoja es absorbida por el parénquima situado debajo; éste actúa, en lo que a captación de luz se refiere, con independencia del parénquima de la otra cara. Esta absorción lumínica por ambas caras quizás explique la elevada cantidad de clorofila presente en las hojas del bosque mediterráneo.

La elevada concentración de clorofila no es una característica exclusiva del bosque mediterráneo. Muchas coníferas tienen en sus acículas concentraciones que superan los 1000 mg m^{-2} cuando se expresan referidas a la proyección plana de la hoja. Sin embargo, estos valores bajan hasta unos niveles normales, en torno a los 400 mg m^{-2}, cuando la concentración de clorofila se refiere a la superficie foliar real, considerando que la sección de las acículas es semicircular. Este paralelismo entre el bosque esclerófilo mediterráneo y los bosques de coníferas no es el único que existe, como veremos al hablar del uso de los nutrientes.

La disposición espacial, aparentemente aleatoria, de las hojas de la encina y de la mayor parte de las plantas del bosque mediterráneo permite una reflexión repetida de la luz dentro de la bóveda, mucho más que cuando las hojas se disponen horizontales, como en el caso del haya; el hecho de captar luz por ambas caras adquiere entonces



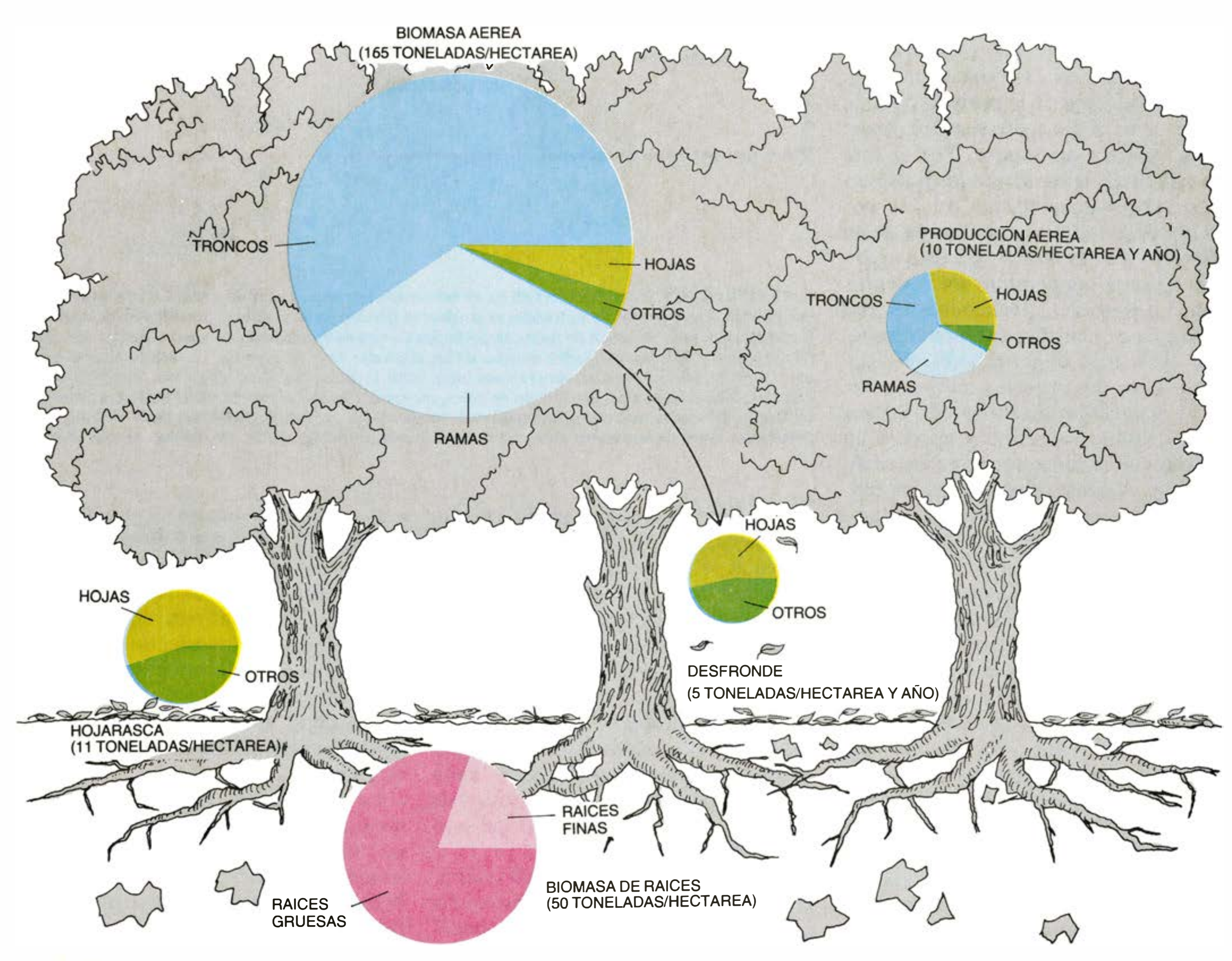
3. CIRCULACION DE NUTRIENTES en un ecosistema terrestre. Las fronteras del ecosistema se indican en rojo. Las entradas de nutrientes al ecosistema (*flechas en azul oscuro*) proceden de la atmósfera (precipitación, sedimentación de polvo, impactación de aerosoles, absorción de gases y fijación biológica de nitrógeno) y de la meteorización química de los minerales de la roca madre. Las salidas de nutrientes del ecosistema (*flechas en azul claro*) tienen lugar hacia la atmósfera (desnitrificación, volatilización de gases, emisión de aerosoles, producción de polen, etcétera) y hacia los arroyos con el agua de avenamiento. Dentro del ecosistema se han distinguido tres componentes: vegetación, mantillo y suelo. Se indican los principales flujos de nutrientes entre estos compartimentos (*flechas rosas*), que forman el ciclo interno.

más importancia. La capacidad de absorber la luz, con esta disposición, disminuye con respecto a los bosques cuyas hojas se sitúan horizontalmente. En el Montseny el encinar tiene ocho metros cuadrados de hojas por cada metro cuadrado de suelo y absorbe aproximadamente la misma luz que el hayedo del mismo macizo montañoso, con tan sólo cuatro metros cuadrados de hojas por metro cuadrado de suelo. Con la distribución aleatoria que presentan las hojas del encinar, el bosque puede mantener mayor número de hojas en el vuelo y acumular así una mayor reserva de nutrientes en las copas. Pero, lógicamente, tal estrategia sólo está justificada en las especies perennifolias como la encina.

La biomasa vegetal es otro de los componentes esenciales de la estructura del bosque. En la parcela experimental del Montseny, con pendiente escasa y un desarrollo particularmente bueno de las encinas, la biomasa arbórea aérea del encinar es de 160 toneladas de peso seco por hectárea (en adelante, t ha^{-1}), de las cuales seis corresponden a hojas y el resto a ramas y troncos. Los estratos arbustivo y herbáceo, muy poco desarrollados en esta parcela, sólo añaden 5 t ha^{-1} a la biomasa del bosque. En Prades, con un bosque de desarrollo medio, la biomasa de la encina es de 100 t ha^{-1}; hay aquí un estrato arbustivo importante, dominado por el madroño y el labiérnago, que suma 15 t ha^{-1}. La biomasa de hojas de encina es de 5,8 t ha^{-1}, y la de otras especies 0,8 t ha^{-1}, resultando una biomasa foliar parecida a la del Montseny. La biomasa foliar de los encinares es pues mayor que la que se encuentra generalmente en bosques caducifolios (1-5 t ha^{-1}), pero menor que la de la mayor parte de los bosques de coníferas (5-20 t ha^{-1}).

Más difícil resulta evaluar la biomasa subterránea. La biomasa de raíces finas (de menos de cinco milímetros de diámetro) varía ampliamente en el tiempo y en el espacio. Muestreos provisionales en el Montseny dan cifras entre 6 y 15 t ha^{-1} en los primeros 40 centímetros del suelo. La excavación de un árbol medio sugiere que la biomasa de raíces gruesas es de unas 40 t ha^{-1} hasta 50 centímetros de profundidad, cifra que, aunque provisional, pone de manifiesto la importancia de la biomasa subterránea del encinar.

La producción neta aérea de las encinas arbóreas en la parcela del Montseny es de 9,3 toneladas de peso seco por hectárea y año; de ellas 5,6 corresponden a madera y corteza, 3,1 a hojas y 0,6 a frutos. Comparando producción y biomasa de hojas, se deduce que el tiempo medio de renovación de éstas es de 2 años. Desconocemos la producción subterránea, aunque dada la rápida renovación de las raíces finas encontrada en otros ecosistemas forestales y la elevada biomasa de raíces gruesas en



4. DISTRIBUCION POR FRACCIONES de la biomasa, la hojarasca, el desfronde y la producción primaria aérea de una parcela experimental en el encinar del Montseny. Los círculos son proporcionales a las cantidades que representan. Las biomasas y la hojarasca son compartimientos, mientras que la producción y el desfronde son flujos. El desfronde está ya incluido en la producción. La biomasa de raíces se evaluó en los primeros 50 centímetros.

el encinar, es de suponer que dicha producción sea considerable.

La comprensión de la economía de los nutrientes va unida al conocimiento de los ciclos biogeoquímicos, tanto en condiciones naturales como tras perturbaciones experimentales. El ciclo de los elementos nutritivos en el bosque resulta de la interacción entre numerosos procesos físicos, químicos y biológicos; constituye uno de los aspectos básicos de su funcionamiento como sistema. Los nutrientes pueden entrar en el ecosistema, salir de él o circular en su seno. La circulación interna se realiza mediante flujos de materia entre las partes del ecosistema, mientras que las entradas y salidas involucran mecanismos de transporte (atmosféricos, hidrológicos y biológicos) que conectan ecosistemas vecinos o lejanos. A través del suelo, de la vegetación y de la fauna, el ecosistema regula en gran medida los flujos de circulación interna y los de salida de nutrientes, pero posee escaso control sobre los flujos de entrada.

La biomasa del bosque contiene cantidades importantes de nutrientes. Destaca el calcio, seguido a cierta distancia por el nitrógeno y el potasio y, a mayor distancia aún, por el magnesio y el fósforo. La parcela del Montseny almacena 1200 kg ha^{-1} de calcio en su biomasa aérea, pese a encontrarse sobre roca silicatada. Los nutrientes contenidos en la vegetación circulan por el ecosistema con tasas rápidas o lentas, según sea la tasa de renovación del órgano vegetal en que se encuentren. Así, la mayor parte de los nutrientes contenidos en la madera quedan secuestrados del ciclo general durante décadas o quizá siglos, mientras que los nutrientes de frutos, hojas y raíces finas circulan activamente entre la vegetación y el suelo. Como es regla general en otras especies de árboles, la madera es el tejido más pobre en nutrientes; las hojas, flores y frutos, los más ricos. El contenido de nutrientes de las hojas varía según su edad y su posición dentro del vuelo, pero siempre son relativamente ricas en nitrógeno y, en menor grado, en fósforo y potasio. Las hojas del encinar, con sólo un cuatro por ciento de la biomasa aérea, almacenan un 20 por ciento del nitrógeno de la vegetación. La corteza de la encina es muy rica en calcio; el 66 por ciento del calcio de la vegetación está contenido en esta fracción, que representa sólo el 17 por ciento de la biomasa aérea del bosque. En términos generales, la mayor parte de los nutrientes se hallan en la madera, porque su gran biomasa compensa las bajas concentraciones: la madera almacena el 79 por ciento del fósforo, el 68 por ciento del potasio y el 48 por ciento del nitrógeno que hay en el total de la biomasa aérea.

Pese a lo que acabamos de exponer, cuando se compara la encina con el conjunto de especies arbóreas de clima templado, resulta evidente que la hoja de encina es pobre en nutrientes; por su contenido en elementos nutritivos se parece a las hojas de las coníferas. Tanto las coníferas como la encina se diferencian de los caducifolios por sus concentraciones foliares muy inferiores, especialmente en nitrógeno y potasio. Recientemente, y sobre todo a partir de los trabajos de R. L. S. Specht en Australia, se ha afirmado la tendencia a pensar que la escasez de nutrientes puede haber desempeñado un papel fundamental en la evolución de los vegetales mediterráneos y, en particular, en el desarrollo de algunas características de sus hojas (abundancia de materiales estructurales como celulosa y lignina y escasez de proteínas) que determinan su esclerofilia y su baja capacidad fotosintética. Si consideramos que la aparición de los climas mediterráneos parece constituir un fenómeno bastante reciente, estas adaptaciones a la escasez de nutrientes habrían resultado ventajosas también para la regulación de la economía hídrica, al producirse la degradación climática con la disminución de las precipitaciones estivales.

Por el contrario, la madera de encina es en general más rica en nutrientes que la de los árboles caducifolios. Si parte de estos nutrientes fuesen movilizables, por lo menos los de las ramas de menor diámetro, la encina dispondría de unas reservas importantes con las que hacer frente a las fluctuaciones en la disponibilidad de los nutrientes del suelo, ocasionadas por la frecuente desecación de los horizontes superficiales. De hecho, los esclerófilos mediterráneos muestran ventaja respecto a los caducifolios de invierno sobre suelos pobres, incluso en zonas donde el clima favorece a los caducifolios. Tomados en conjunto, el bajo contenido en nutrientes de las hojas de encina (indicador de unos requerimientos nutritivos también bajos) y la relativa riqueza en nutrientes de las ramas parecen configurar un síndrome adaptativo basado en una economía frugal de los nutrientes y en un cierto control interno que favorecerían la colonización de medios pobres y fluctuantes.

Tarde o temprano, los nutrientes contenidos en la vegetación han de volver al suelo del bosque. Las principales vías de retorno de nutrientes son el desfronde (caída de hojas, frutos, ramas, etcétera), la descomposición de las raíces finas y la lixiviación de las plantas producida por el agua de la lluvia.

El desfronde es de vital importancia en el funcionamiento del bosque, no sólo como vía de retorno de nutrientes sino porque constituye, junto con la descomposición de las raíces finas, el principal flujo de materia orgánica al suelo del bosque, que es a su vez la puerta de entrada de la energía a la red trófica de los organismos descomponedores y detritívoros (hongos, bacterias, microartrópodos y otros). Además, la magnitud del desfronde guarda una correlación estrecha con la producción primaria del bosque. Así, el desfronde anual es de 5,3 t ha^{-1} en el encinar del Montseny y de 3,3 t ha^{-1} en Prades, diferencia que probablemente refleja la menor productividad de este último encinar, debida a la menor cuantía de las precipitaciones.

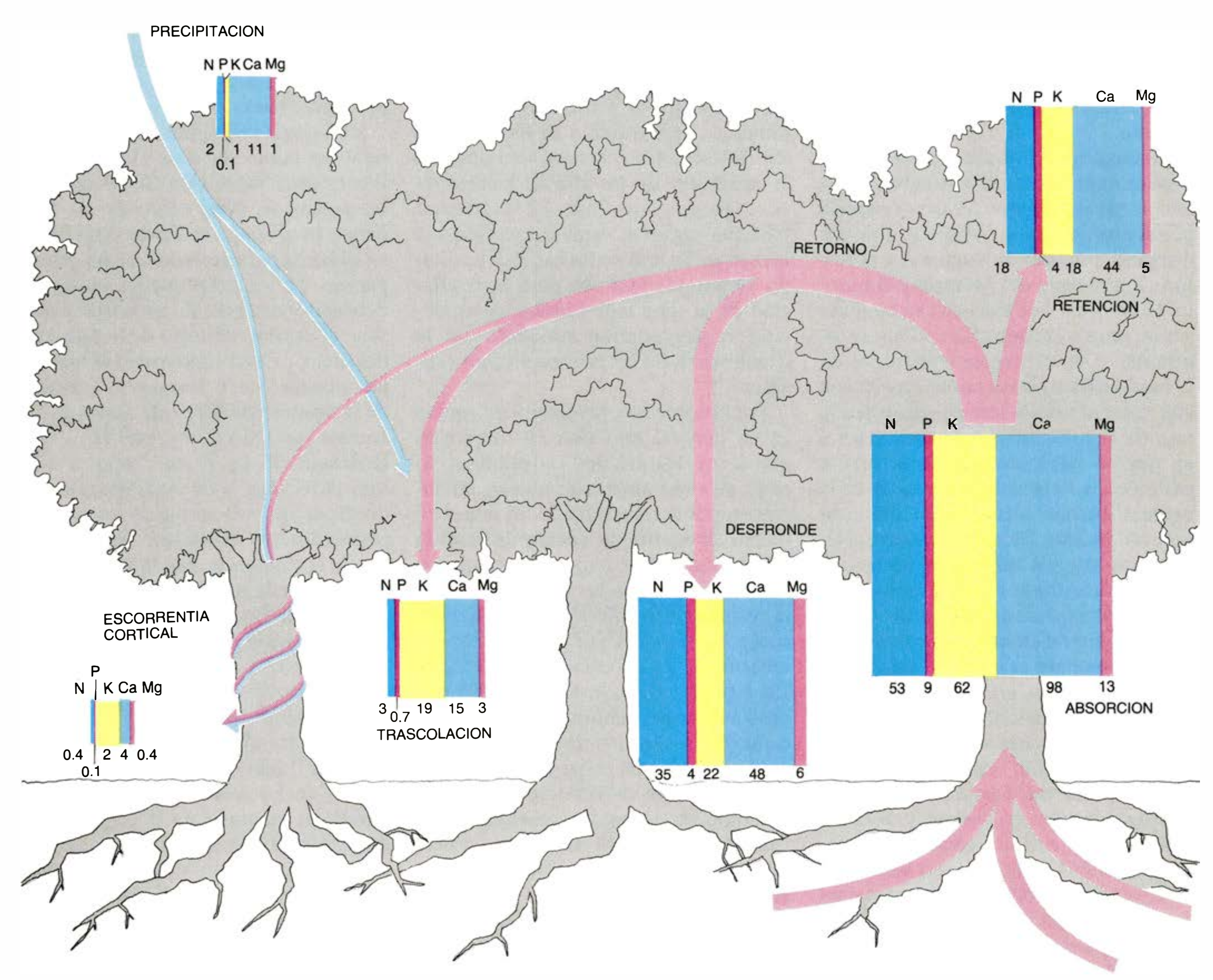
El flujo de nutrientes en el desfronde tiene un papel clave en el ecosistema. Previa descomposición de los materiales orgánicos, estos nutrientes se reincorporan al conjunto de los disponibles en el suelo para su reutilización por las plantas. Se ha dicho que los nutrientes del desfronde, más los de la lixiviación, son “el capital rotatorio de la fertilidad del suelo”. Centrándonos en el nitrógeno, elemento con frecuencia limitante de la producción forestal, vemos que el encinar del Montseny recicla con el desfronde 35 kg N ha^{-1} $año^{-1}$, cifra muy parecida a la de otros bosques mediterráneos y a la media de los bosques templados de coníferas. En general, puede considerarse que la magnitud de los aportes de nitrógeno en el desfronde es indicativa de la disponibilidad de nitrógeno en el ecosistema. Así, el desfronde de los bosques templados caducifolios, normalmente más eutróficos que los de coníferas y los mediterráneos, aporta al suelo una media de 51 kg N ha^{-1} $año^{-1}$.

Uno de los mecanismos de que disponen las plantas para la conservación de nutrientes es la retranslocación: la retirada de nutrientes de los órganos viejos, en particular las hojas, hacia las hojas jóvenes, las flores o los frutos. Por ejemplo, las hojas verdes de encina contienen a los dos años de edad una media de 13 mg N g^{-1} (peso seco), cantidad que disminuye a 6 mg g^{-1} en las hojas recién caídas. Parte de esta pérdida podría deberse a la lixiviación de nitrógeno orgánico o a la volatilización

de amoníaco en las hojas senescentes, pero la mayor parte debe ser retranslocada antes de la caída de la hoja hacia órganos que estén en crecimiento activo. Proporciones significativas de fósforo y de magnesio (pero no de calcio ni, sorprendentemente, de potasio) son también retiradas de las hojas de encina antes de su caída.

Otra de las vías de retorno de nutrientes es la lixiviación por el agua de lluvia. Esta puede llegar al suelo del bosque cayendo a través de las copas (trascolación) o resbalando por los troncos (escorrentía cortical). En ambos casos, el agua se enriquece en la mayor parte de los nutrientes, debido al lavado de las copas y de los troncos. Este lavado es una mezcla prácticamente inseparable de los nutrientes que han entrado en el ecosistema procedentes de la atmósfera por deposición seca y de nutrientes que, habiendo sido absorbidos del suelo por las plantas, circulan internamente por el ecosistema gracias a la lixiviación de las hojas y la corteza. A pesar de esta incertidumbre sobre el origen de los nutrientes lavados de las copas y los troncos, puede afirmarse que la trascolación y la escorrentía cortical constituyen una vía muy importante de retorno de potasio, ya que los elevados aportes de este elemento no pueden atribuirse, más que en una pequeña parte, a la deposición seca. Los aportes de potasio en la trascolación y la escorrentía cortical tienen una marcada pauta estacional en el encinar, con un máximo destacado en el momento de mayor caída de la hoja entre mayo y agosto. El retorno de fósforo y magnesio está bastante correlacionado con el de potasio, lo que induce a pensar que la lixiviación sería también responsable de una buena parte del lavado de ambos elementos. En Prades, se observa una retención de nitrato por parte de las copas y de los troncos, hecho que no es infrecuente en los bosques templados, si bien no se da en la trascolación del Montseny. Dado que las cutículas de las hojas de encina son gruesas, y como la absorción de nitrato se observa también en los troncos, es probable que esta desaparición de nitrato no se deba a una absorción foliar, sino a la actividad de la microflora epifítica.

Los materiales del desfronde se acumulan temporalmente sobre el suelo del bosque, formando el mantillo. En el Montseny, la biomasa del mantillo es de 11 t ha^{-1} y, en Prades, de 14, cifras pequeñas en el contexto de los bosques templados que indican una activa des-

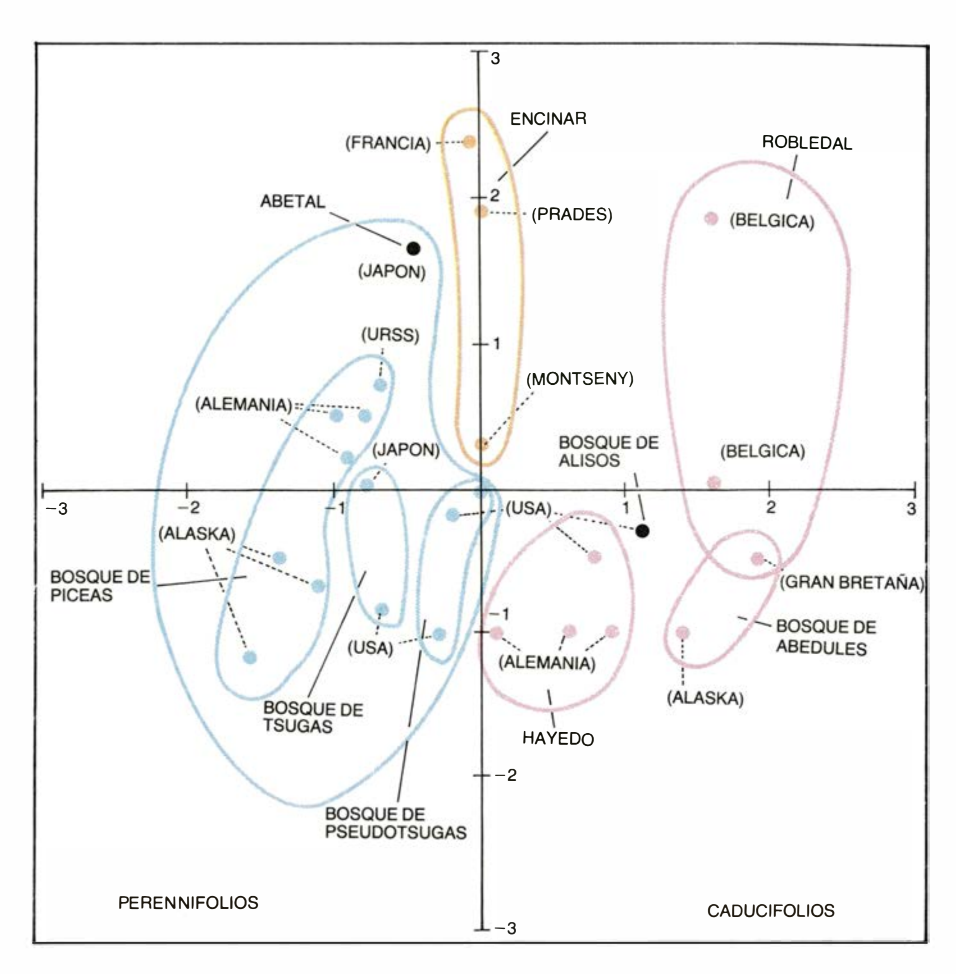


5. CICLOS BIOGEOQUIMICOS de cinco elementos en el encinar del Montseny. Las cifras son kilogramos por hectárea y por año. Las flechas azules representan entradas al ecosistema; las rojas, el ciclo interno. La retención la constituyen los nutrientes anualmente acumulados en la madera y corteza nuevas. El retorno de nutrientes es la suma del desfronde más lixiviación. La absorción es la suma del retorno y la retención. Para calcular el retorno se ha supuesto que los nutrientes lavados de los árboles por la trascolación (lluvia dentro del bosque) y por la escorrentía cortical (agua que resbala por los troncos) proceden de la lixiviación y no de la deposición seca. Los flujos en la escorrentía cortical han sido estimados a partir de los datos de Prades.

composición de la materia orgánica en los suelos de estos encinares, cuya temperatura media anual es bastante alta. El encinar de Prades, con un desfronde muy inferior, acumula más mantillo que el del Montseny, debido a un retardo de la descomposición en un clima menos lluvioso.

El mantillo almacena cantidades variables de los distintos nutrientes, según su movilidad. La relación entre la cantidad de nutrientes en el mantillo y los aportes por el desfronde mide el tiempo de renovación de los elementos. En el Montseny, estos tiempos en años son: materia orgánica, 2,0; potasio, 1,0; sodio, 1,3; magnesio, 2,0; fósforo, 2,3; nitrógeno, 3,1; y calcio, 3,2. Estas cifras nos indican que el potasio se pierde de los materiales del mantillo a una velocidad mucho mayor que la materia orgánica; el magnesio se libera a la misma velocidad a la que se mineraliza dicha materia, mientras que el nitrógeno y el calcio se liberan más lentamente y, por tanto, los materiales del mantillo estarán enriquecidos en estos dos nutrientes respecto a los materiales del desfronde. Los anteriores tiempos de renovación son considerablemente menores que los habituales en bosques caducifolios templados y mucho menores que los que se dan en bosques templados de coníferas. Por ejemplo, en estos últimos, el tiempo medio de residencia del nitrógeno en el mantillo es de 18 años.

Vista la importancia de la escasez de nutrientes en la ecología de la vegetación mediterránea, vale la pena examinar con qué eficacia la encina utiliza el nitrógeno. Un índice de esta eficacia nos lo ofrece el cociente entre la producción neta del bosque y la absorción de nitrógeno. El encinar del Montseny produce 181 kilogramos de materia orgánica por cada kilogramo de nitrógeno absorbido. Los bosques templados de coníferas producen una media de 184 kilogramos por cada kilogramo de nitrógeno, mientras que los caducifolios templados producen en promedio sólo 138 kilogramos por cada uno de nitrógeno. Una vez más, las encinas se alinean con las coníferas y difieren de los caducifolios; ello indica un uso eficaz del nitrógeno y una adaptación a medios cuya disponibilidad de nutrientes es limitada.



6. ANALISIS MULTIVARIANTE de 25 bosques de todo el mundo. Dicho conjunto se ha definido por 30 variables que incluyen biomasas y producciones de hojas y ramas y sus contenidos en nutrientes. La ordenación en componentes principales da un primer eje en que dominan las variables relativas a las hojas; este eje separa los bosques caducifolios (*en rojo*), con mayor contenido en nutrientes, de los bosques de coníferas (*en azul*), mientras que los encinares (*en amarillo*) aparecen entre ambos grupos pero más próximos a estas últimas. El segundo eje, dominado por las variables relativas a los nutrientes en ramas, ordena los bosques de suerte tal que sugiere un gradiente de eutrofía en el suelo. Las afinidades taxonómicas quedan bien reflejadas en la agrupación de los puntos en el plano definido por los dos componentes.

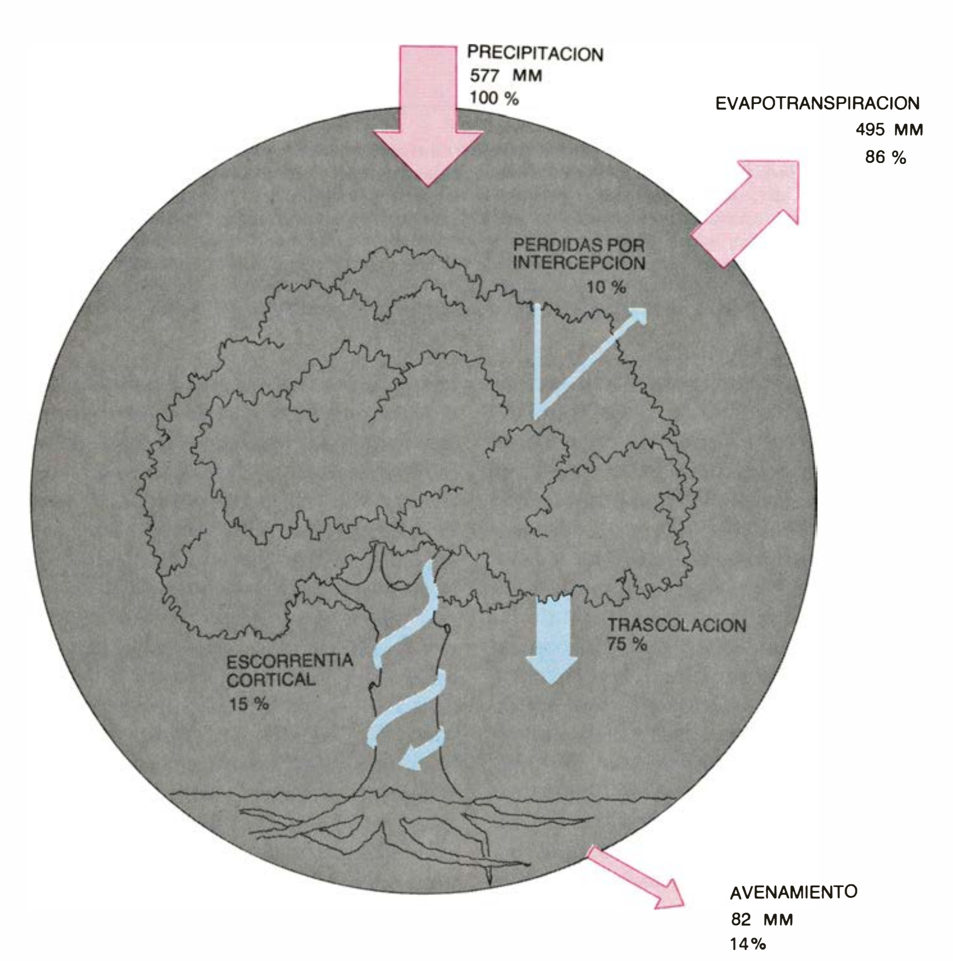
La comprensión de los complejos mecanismos que aseguran el funcionamiento del ecosistema requiere de estudios detallados que incluyen no sólo los temas que hemos tratado hasta aquí, sino también otros muchos: el seguimiento detallado de la fenología del bosque, la distribución de recursos en la estructura, el desarrollo estacional de los procesos de crecimiento, el papel de los microartrópodos en la degradación de los restos, el transporte de nutrientes realizado en el suelo por los micelios de los hongos, la relación de las tasas de actividad en el suelo con las condiciones de humedad y temperatura y muchos otros temas que no hemos abordado son de importancia fundamental en la vida del bosque. No obstante, y pese a la necesidad de avanzar en todos estos conocimientos, la problemática de la gestión del bosque mediterráneo se ha hecho acuciante e invita a prestar atención prioritaria a algunos aspectos muy generales del metabolismo global del mismo. En un artículo publicado en esta misma revista, F. H. Bormann y G. E. Likens planteaban la utilidad de considerar ecosistemas delimitados por pequeñas cuencas hidrológicas, con vegetación uniforme y sobre sustrato impermeable, en las que es posible establecer con precisión el balance de entradas y salidas de agua y nutrientes. En este método, el bosque es considerado globalmente como una caja negra. Disponiendo de sistemas de cuencas similares, dotadas de los equipos adecuados (pluviógrafos, registradores de caudal, muestreadores de agua de lluvia y de escorrentía para análisis), es posible realizar experimentos controlados de los efectos de diversos métodos de gestión o impactos, por ejemplo, limpieza del sotobosque, quema controlada o incendios, extracción maderera de distinta intensidad, etc., sobre el balance global de agua y nutrientes del ecosistema.

Este tipo de aproximación no se había intentado nunca en el bosque mediterráneo. Nuestro programa conjunto nos ha permitido dar unos primeros pasos para conocer el funcionamiento normal del encinar no sometido a inter-

vención y el papel que tienen las tormentas en este funcionamiento.

Un primer aspecto de interés en el uso de pequeñas cuencas aforadas es la posibilidad de calcular balances hídricos. Si se supone que la impermeabilidad de la roca madre no permite pérdidas hacia acuíferos profundos, en un período de tiempo comprendido entre dos momentos en que se den similares condiciones hídricas del suelo se puede afirmar que las entradas de agua por precipitación deben ser equivalentes a las salidas por avenamiento y por evapotranspiración. En las cuencas experimentales, se miden de modo continuo la precipitación y el avenamiento. Su diferencia proporciona una buena estima de la evapotranspiración.

En los treinta meses en que el Barranc de l'Avic (Prades) ha tenido control hidrológico, el arroyo se secó en dos ocasiones, separadas por un intervalo de un año. En este período se puede suponer que los estados hídricos inicial y final del suelo son semejantes; resulta entonces posible el cálculo de un balance de agua. La entrada por precipitación durante estos doce meses fue de 577 milímetros y la salida por avenamiento 82 milímetros, es decir, sólo un 14 por ciento de las entradas. Los restantes 485 milímetros, que corresponden al 86 por ciento de las entradas, representan la salida por evapotranspiración, cifra que se muestra coherente con los datos de transpiración del encinar que se han obtenido por otros métodos. El porcentaje de agua que pasa a la atmófera en la evapotranspiración depende en gran medida del régimen pluviotérmico del lugar y de las características de la cubierta vegetal. En una serie de casi 40 cuencas experimentales, sobre todo de países templado-húmedos, la mayoría de los valores porcentuales de evapotranspiración son menores del 50 por ciento, pero también se han citado unos pocos casos superiores al 80 por ciento en bosques de coníferas y en el chaparral, aunque en algunos en el flujo se incluyen pérdidas de agua por el subsuelo.



7. CIRCULACION DEL AGUA en el encinar de Prades. Las entradas y salidas del ecosistema se expresan en milímetros anuales. Del total caído, en torno a un 10 por ciento se evapora directamente desde las copas, antes de llegar al suelo (las llamadas pérdidas por intercepción). La escorrentía cortical alcanza una proporción alta, probablemente debido a la gran densidad de pies en este encinar. La elevada evapotranspiración, que alcanza un 86 por ciento del agua recibida, indica el alto grado de control que tiene el ecosistema no sólo sobre el flujo del agua sino también sobre los nutrientes que ésta exporta del sistema.

Con porcentajes tan elevados de retorno de agua a la atmósfera y cuando, como ocurre en el encinar, éstos se deben sobre todo a la transpiración, nos hallamos ante un control biológico importante del balance hídrico del ecosistema, lo que naturalmente afecta también a las exportaciones de sustancias en forma disuelta y particulada, reduciéndolas.

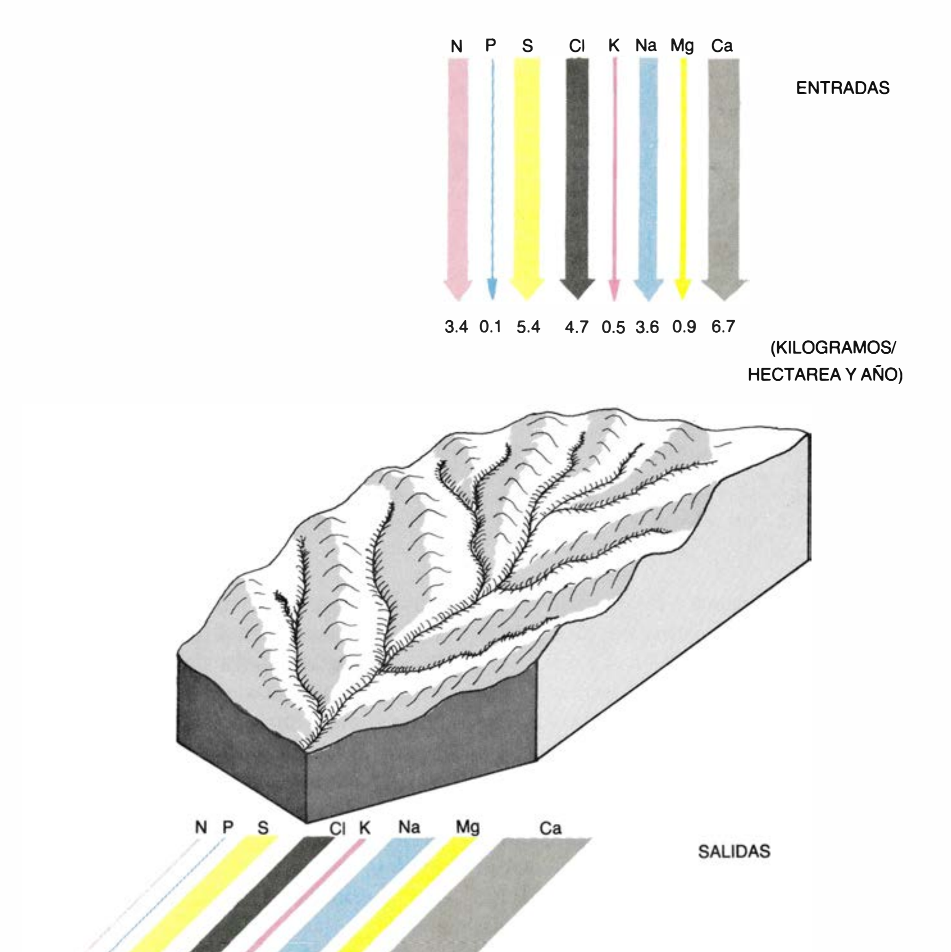
El agua que avena la cuenca lleva cierta cantidad de sustancias disueltas y arrastra además una serie de partículas orgánicas e inorgánicas de diferente tamaño. Con una trampa de sedimentos, con una red de malla fina en el aliviadero y con el filtrado de muestras del agua saliente a través de filtros de 0,45 micrometros, se consigue una buena recolección del material particulado que abandona la cuenca experimental. Sorprendentemente, las primeras estimas de este flujo en Prades son sólo del orden de 0,8 kg ha^{-1} $año^{-1}$, mientras que la exportación en forma disuelta alcanza los 47 kg ha^{-1} $año^{-1}$. Como contraste comparativo, las primeras cifras de erosión de suelo calculadas dentro del proyecto LUCDEME (Lucha Contra la Desertización en el Mediterráneo) en varias cuencas hidrográficas del sureste de España oscilan entre 100 y 940.000 kg ha^{-1} $año^{-1}$; se calcula que un 76 por ciento del territorio abarcado por el proyecto pierde anualmente más de 10.000 kg ha^{-1}. Desde el punto de vista de la metodología, resulta más válida la comparación con el valor de 33 kg ha^{-1} $año^{-1}$ de materia particulada total obtenida como media de ocho años en una cuenca con bosque caducifolio en Hubbard Brook (New Hampshire) que, no obstante, tiene un caudal anual diez veces mayor que el de l'Avic.

Desde luego, el menguado avenamiento de la cuenca del encinar está muy condicionado por la transpiración del bosque; además, la capacidad de arrastre del agua se ve también limitada por ser el recorrido subálveo en muchos tramos, así como por la estabilidad de los acúmulos de restos orgánicos en el cauce del arroyo que actúan a modo de filtros.

El caudal de avenamiento mensual permite una buena predicción del peso total de materia disuelta en él, ya que existe una buena relación lineal entre ambos. Con la materia particulada total, la relación es exponencial y peor el ajuste estadístico. Curiosamente, la exportación de materia particulada inorgánica es predecible, incluso linealmente, a partir del caudal, cosa que no

sucede con la orgánica. Esto parece sugerir que la movilización de restos orgánicos en cantidad excepcional se daría por desestabilización de los acúmulos existentes en el lecho del arroyo, que se produciría bruscamente a partir de una determinada energía del agua, fenómeno que afectaría menos a los materiales inorgánicos.

Es muy probable que estos arrastres inusuales se den en lapsos de tiempo muy cortos. Bormann y sus colaboradores han estimado para una cuenca de Hubbarb Brook que el 86 por ciento de la materia particulada se pierde en sólo 1,6 por ciento del tiempo y con el 23 por ciento del volumen de agua y que el 16 por ciento lo hace en sólo el 0,0025 por ciento de tiempo y el 0,2 por ciento del caudal. En l'Avic el 39 por ciento de la exportación anual en forma particulada se produce como máximo en el 3 por ciento del tiempo y con el 13 por ciento del caudal. Este carácter episódico e imprevisible de la salida de materiales particulados tiene especial interés por ser la vía privilegiada de exportación de fósforo y hierro.



8. BALANCES DE NUTRIENTES en la cuenca de encinar del Barranc de l'Avic (Prades). Se representa las entradas con el agua de lluvia (más la sedimentación de polvo) y las salidas en forma disuelta con el agua del arroyo. Se considera sólo las fracciones inorgánicas de los distintos elementos. La cuenca actúa como sumidero de nitrógeno inorgánico y como fuente de calcio, sodio y magnesio, que procederían de la meteorización química de los silicatos del suelo o de las pizarras que forman la roca madre.

Con el control de las entradas y salidas de nutrientes en el ecosistema cuenca, es posible el cálculo de balances. E. P. Odum supuso inicialmente que las pérdidas serían mínimas en los ecosistemas más maduros, de los que drenarían las aguas más puras. Pero esta hipótesis no parece confirmarse. Si analizamos los compartimentos y flujos que intervienen en uno de estos balances (roca, suelo, mantillo y vegetación entre los primeros y deposición global, meteorización y avenamiento, entre los segundos), admitiendo que en un bosque en crecimiento se acumulan elementos en la biomasa y que el suelo se halla próximo al estado estacionario, una parte de las entradas por vía atmosférica o desde el substrato geológico se incorporarían al ecosistema. Por el contrario, en un bosque que casi no crezca (cuya respiración se aproxime a su producción primaria), contrariamente a lo supuesto por Odum será previsible un excedente de nutrientes en virtud de esas entradas. Lógicamente, el comportamiento de los distintos elementos químicos no es el mismo, sino que depende del factor de concentración con que pasen a formar parte de los vegetales.

Por consiguiente, los balances de diferentes elementos nos sirven de indicadores del estado funcional del ecosistema, pero siempre que se tenga en cuenta que las entradas atmosféricas, especialmente las debidas a la precipitación, poseen una importante componente estocástica y siempre que se considere las reservas minerales de la roca y la velocidad a que se liberan.

Un problema adicional en la interpretación de balances se debe a la capacidad de control del bosque sobre las salidas, que es función del tiempo de permanencia del agua en el sistema. El carácter irregular de la distribución de las precipitaciones, propio del clima mediterráneo, y la gran intensidad de la lluvia en determinadas tormentas, especialmente en verano, serían factores negativos para el control de las exportaciones por parte del encinar. Por el contrario, el alto porcentaje de agua que constituye la evapotranspiración y el hecho de que las precipitaciones de mayor intensidad coincidan con los momentos de máximo déficit hídrico en el suelo serían elementos que favorecerían una estrategia conservadora por parte del bosque esclerófilo mediterráneo. Un buen ejemplo de proceso disruptor lo constituye una tormenta de 26 milímetros ocurrida en una cuenca del Montseny en agosto de 1983, que fue seguida con una muestra cada siete minutos y medio y en la que, aunque sólo drenó el 2 por ciento de la precipitación incidente, se vio un pulso de nitratos con concentraciones hasta 50 veces superiores a las iniciales. A modo de contraste y para ejemplificar una situación de retardo, mencionemos la tormenta de 76 milímetros de precipitación que tuvo lugar en l'Avic el 27 de agosto de 1982. La lluvia se produjo después de 40 días de haberse secado el arroyo y originó un primer pico que representa sólo el 0,01 por ciento de la precipitación; se produjo un nuevo y

ELEMENTO	BALANCE NETO (KILOGRAMOS POR HECTAREA Y AÑO)		CICLO INTERNO EN L'AVIC (KILOGRAMOS POR HECTAREA Y AÑO)
	L'AVIC	MEDIA BIBLIOGRAFICA	
N	3,2	5,4	17,9
P	0,1	0,1	1,4
S	−2,5	−5,4	9,8
Cl	−2,8	0,3	4,8
K	−0,7	−2,3	31,9
Na	−5,6	−11,0	3,3
Mg	−4,9	−9,1	4,2
Ca	−22,6	−24,5	5,0

9. BALANCES NETOS de ocho elementos químicos en cuencas forestales. Se calculan por diferencia entre los flujos de entrada con el agua de lluvia y los flujos de salida con el avenamiento. Los balances netos positivos indican que el elemento se está acumulando en el ecosistema; los negativos, un empobrecimiento del sistema o la existencia de una fuente interna a la cuenca, como es la meteorización. Se comparan los resultados de l'Avic con una media bibliográfica recopilada por G. E. Likens y colaboradores. Como contraste, se incluye la magnitud del retorno de nutrientes desde la vegetación al suelo, que mide la intensidad del ciclo interno al ecosistema. (Solamente se incluye la fracción inorgánica disuelta.)

breve agostamiento y luego otra subida del caudal, a las 70 horas de iniciada la tormenta.

A partir del balance hídrico y de las concentraciones químicas ponderadas por volumen en el agua de lluvia y del arroyo, se han obtenido los flujos anuales de entrada y salida de nutrientes en la cuenca de l'Avic. Por diferencia entre las entradas y las salidas se obtiene el balance neto de cada elemento. Cuando las entradas superan las salidas, el elemento en cuestión está acumulándose en el ecosistema y su balance neto es positivo. Tal es el caso del nitrógeno, o por lo menos del nitrógeno inorgánico, tanto en l'Avic como en las demás cuencas estudiadas por otros autores. Es decir, los bosques en crecimiento actúan de sumideros de nitrógeno; evitan así que este nitrógeno acabe llegando a los sistemas acuáticos situados aguas abajo y agrave los problemas de eutrofización. Por otra parte, los aportes atmosféricos de nitrógeno, en parte debidos a la contaminación del aire, constituyen un subsidio para el ecosistema forestal, en forma de fertilización crónica de bajo nivel.

A diferencia del nitrógeno, la mayor parte de los elementos tienen balances netos negativos. Esto significa que hay una fuente interna a la cuenca (probablemente la meteorización) o que el sistema se está empobreciendo en dichos elementos. En l'Avic, calcio, magnesio, sodio, potasio, azufre y cloro tienen balances netos negativos. En particular, hay una pérdida neta importante de los tres primeros elementos, debida a la meteorización química de los silicatos. Se observa que en esa zona de Prades, con 80 milímetros de avenamiento anual, la salida neta de calcio es similar, y las de magnesio y sodio del orden de la mitad, al valor medio de un conjunto de cuencas recopiladas por otros autores, que avenan en promedio unos 800 milímetros anuales. Es decir, la exportación neta de estos tres elementos por unidad de flujo de agua que abandona la cuenca es muy superior en l'Avic. Esto indicaría, dado que los esquistos no son aquí especialmente atacables, que la meteorización se ve favorecida en esta cuenca mediterránea por la elevada temperatura media y, quizá, por la alternancia de períodos húmedos y secos.

A partir del balance del sodio, se estima para las pizarras de l'Avic una tasa de meteorización anual de 613 kg ha^{-1}, lo que representa la descomposición de 2,4 milímetros de roca cada cien años. Con esta tasa de meteorización se liberan casi 20 kilogramos de potasio, 12 de azufre, seis de sodio, cuatro de magnesio, dos de calcio y alrededor de 0,25 kg ha^{-1} $año^{-1}$ de fósforo, que constituyen, excepto en el caso del calcio, cantidades de importancia si se comparan con las cifras de balances y ciclo interno.

Merece algún comentario la detección de lluvias ácidas en nuestras estaciones. Los valores medios del *p*H están entre 4,6 y 4,8; en ciertas tormentas se hallan valores inferiores a 4. Creemos que ésta es la primera constatación del fenómeno de las lluvias ácidas en zonas no urbanas de nuestro país. No obstante, en el Montseny sólo el 35 por ciento de los hidrogeniones libres (H^+) presentes en la deposición global alcanza el suelo con la trascolación, quedando el resto de la acidez neutralizado por la deposición seca acumulada en la vegetación, por intercambio catiónico en las superficies vegetales y espacios intercelulares de las hojas o por lixiviación de sales de ácidos débiles. El contacto con los suelos de las cuencas silicatadas, tanto en Prades como en el Montseny, acaba de neutralizar la acidez, probablemente por la meteorización química de los aluminosilicatos, que consume un equivalente de H^+ por cada equivalente de cationes metálicos liberados a la solución del suelo. Quizás en el estado actual haya una aceleración de la meteorización por las lluvias ácidas, sin pérdidas sensibles en el suelo, pero el intercambio catiónico en éste podría concluir, a la larga, a su empobrecimiento si la deposición de ácido aumentase.

Desde el punto de vista práctico, son altamente prometedoras las posibilidades que el uso de cuencas controladas abre para la experimentación de diversos tipos de gestión del bosque o de los efectos de diferentes clases de aportes forzados. La detección y el seguimiento de los cambios que en los flujos de salida de agua y de materia disuelta y particulada pueden producirse como consecuencia de talas de distinta intensidad, incendios más o menos controlados, diferentes tipos de abonados, limpias manuales de sotobosque, uso de herbicidas, deposiciones ácidas y otras muchas acciones como aterrazamiento, apertura de cauces y otras, pueden ser de gran interés para evaluar los efectos y poder determinar la capacidad de asimilación por el ecosistema en conjunto.

El precedente más conocido de este aspecto experimental lo constituye la tala rasa de toda una cuenca de Hubbard Brook, sin extracción alguna y con un tratamiento de biocidas durante tres años. Los resultados de este drástico experimento han sido del mayor interés en el descubrimiento de los mecanismos de recuperación de la capacidad de control por parte de la vegetación regenerada. La gran ventaja del dispositivo experimental de Hubbard Brook, hoy adoptado en otros centros de investigación, consiste en disponer de una amplia batería de cuencas convenientemente aforadas y de parecidas características en las que es posible realizar, de modo sistemático, el estudio de la respuesta "metabólica" del ecosistema frente a diversos tipos de perturbaciones naturales o debidas a la acción del hombre. Nuestra intención es completar las actuales instalaciones de Prades y del Montseny en este sentido y, si es posible, extenderlas a otros ecosistemas forestales o arbustivos mediterráneos, ya que el método derivado de la aplicación del concepto de ecosistema-cuenca es, sin duda, una de las más potentes vías de aproximación de que disponemos para la comprensión del funcionamiento de los ecosistemas terrestres.

Estructura y evolución de los granos interestelares

Las partículas de polvo cósmico ofrecen una compleja estructura de núcleo y envoltura. En la envoltura orgánica hay indicios que hablan de procesos químicos complicados e incluso violentos

J. Mayo Greenberg

El espacio que media entre las estrellas está lleno de materia. La mayor parte de ella consiste en hidrógeno y helio, que, bajo las condiciones interestelares, no puede formar cuerpos sólidos. Incluso a tres grados Kelvin, el universo está sencillamente demasiado caliente para que puedan condensarse. Pese a ello, hay cuerpos sólidos: unas diminutas partículas congeladas, los llamados granos interestelares. Se forman éstos a partir de elementos químicos más pesados, sintetizados por fusión termonuclear en estrellas y supernovas. De esos elementos, los más abundantes son el oxígeno, el carbono y el nitrógeno; integran el colectivo de los elementos orgánicos. Les siguen en abundancia el magnesio, el silicio y el hierro. Los granos interestelares están compuestos por esos seis elementos condensables, junto con el hidrógeno que capturan.

Como ocurre en otras ramas de la astronomía, el estudio de los granos interestelares se ve dificultado por la inaccesibilidad de la materia a investigar. No se ha podido recoger muestras para su examen en el laboratorio. La única fuente directa de información es la radiación electromagnética que alcanza la Tierra procedente de las estrellas después de atravesar regiones del cielo que contienen polvo interestelar. Observando las longitudes de onda difundidas y absorbidas por los granos, se infieren algunos datos acerca de sus propiedades físicas y composición química. Ultimamente, otro método de estudio ha suministrado valiosas indicaciones: ensayos de laboratorio que han simulado la interacción entre gases, sólidos y radiación en el espacio interestelar. Se pretende crear sustancias que puedan explicar plausiblemente las observaciones astronómicas.

De los trabajos astronómicos y experimentales se ha deducido que los granos interestelares poseen una notable complejidad. No son, ciertamente, meros aglomerados amorfos e inertes de polvo cósmico. Un grano típico tiene una estructura interna definida, con un núcleo constituido principalmente por silicatos (materia de tipo rocoso) y una envoltura de materia orgánica más volátil. En la envoltura, donde se aglomera la materia orgánica, se desarrollan procesos químicos complicados, que quizás incluyan la detonación ocasional de componentes inestables de la mezcla. Los granos se caracterizan, además, por un interesante ciclo de vida. Nacen como "simientes" de silicatos expulsadas de estrellas viejas; después de una larga evolución física y química, acaban su vida absorbidos por estrellas en proceso de formación.

A pesar de la presencia de gas y granos, el espacio interestelar es un vacío mayor que cualquier otro logrado en la Tierra. La densidad media de materia interestelar en la Vía Láctea es de un átomo de hidrógeno por centímetro cúbico. Pero la distribución de la materia está muy lejos de ser uniforme. A gran escala, el gas de una galaxia espiral, tal como la Vía Láctea, se concentra en el plano central del disco galáctico. Parece especialmente denso a lo largo de los bordes internos de los brazos espirales. A escala local, el gas presenta también una distribución irregular: gran parte se ha aglomerado en nubes.

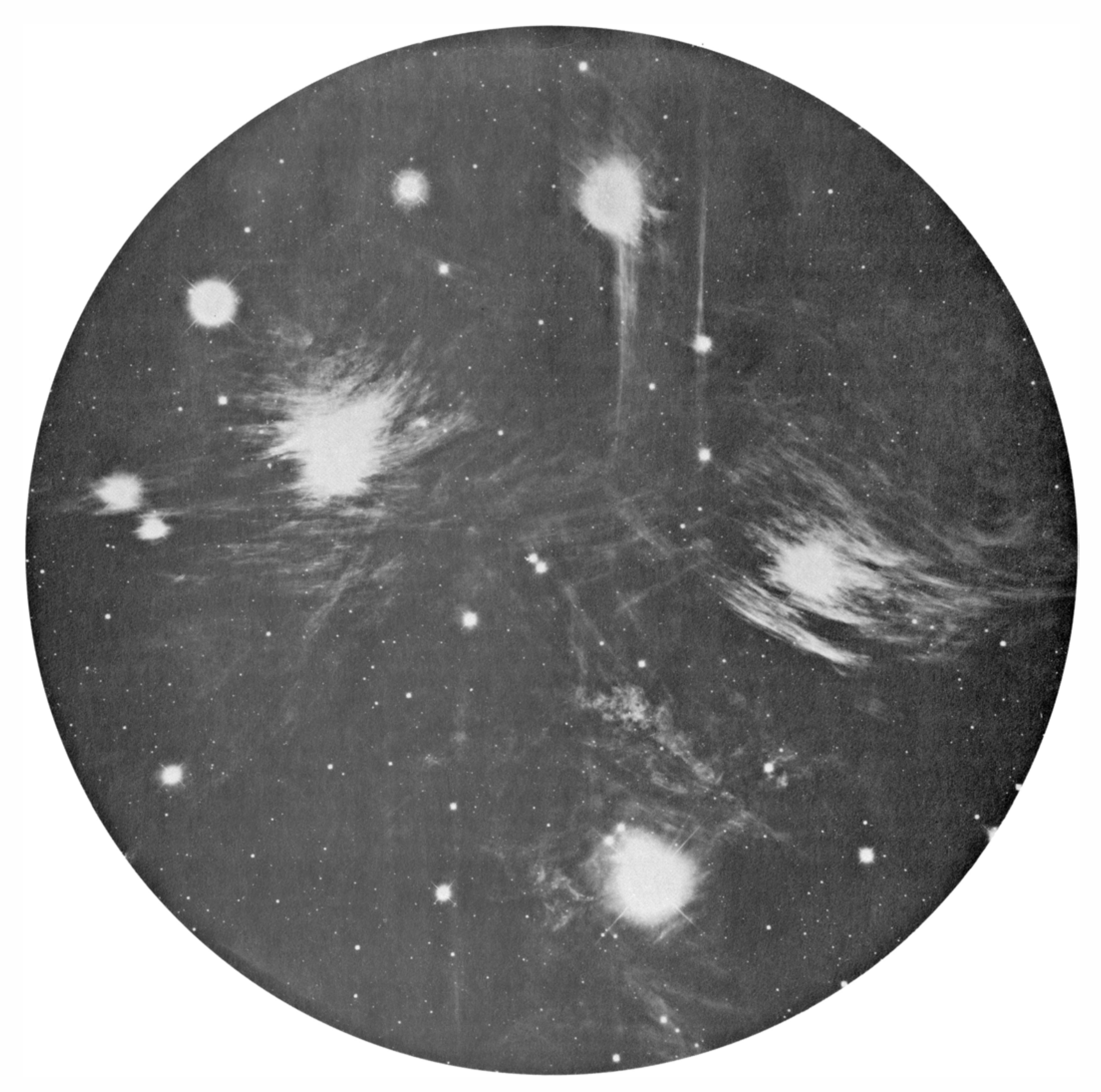
Se distingue entre dos clases de nubes: difusas y moleculares. Las difusas son regiones de baja densidad donde el hidrógeno y otros gases se presentan sólo en forma de átomos aislados. La densidad puede ser inferior a 20 átomos por centímetro cúbico. Las nubes moleculares constituyen regiones de más densidad, donde la mayor parte de los átomos de hidrógeno se han combinado en pares para formar moléculas de hidrógeno diatómico (H_2). Se ha descubierto también moléculas más complejas. En una nube molecular, la densidad alcanza hasta un millón de átomos por centímetro cúbico. Ahora bien, hasta las nubes moleculares más densas resultan de una levedad extremada si las comparamos con la Tierra. Pese a ello, la masa del gas interestelar de la Vía Láctea viene a ser de un diez por ciento de la masa de las estrellas en dicha galaxia.

La distribución de granos interestelares parece corresponder con la del gas. Sin embargo, lo que delata la presencia de granos es su manera de difundir la luz de las estrellas. La interacción entre las partículas y la luz depende fundamentalmente del tamaño de las partículas y de la longitud de onda de la luz. Pensemos en las partículas de polvo de tiza. Son mayores que cualquier longitud de onda de la luz visible, de modo que difunden igualmente todas esas longitudes de onda. La luz que atraviesa una suspensión de polvo de tiza queda, por tanto, atenuada, pero no cambia de color. No así las partículas del tamaño de una longitud de onda determinada, que dispersan dicha longitud más eficazmente. Las partículas cuyo tamaño sea algo menor que una longitud de onda típica de la luz visible difundirán la luz azul más que la roja, con lo que suprimirán esa parte del espectro del haz de luz que llega al observador.

La luz de las estrellas que atraviesa el medio interestelar queda, pues, enrojecida según la pauta de "extinción" que producen los granos interestelares. Esta pauta se ha medido a lo largo de un dominio que comprende desde las

grandes longitudes de onda del infrarojo lejano hasta las muy pequeñas del ultravioleta. Al principio, la extinción aumenta suavemente: a medida que la longitud de onda disminuye desde el infrarrojo, pasando por el visible y penetrando en el ultravioleta próximo, la cantidad de radiación eliminada por difusión aumenta de un modo constante. Después, la curva se hace más empinada, alcanzando un máximo a una longitud de onda de unas 2200 unidades angström. Luego, desciende hasta un mínimo relativo hacia los 1700 angström, para volver de nuevo a aumentar en el ultravioleta lejano.

La interpretación de esta configuración ha sido muy controvertida. Pero el examen minucioso de la pauta registrada por el satélite Explorador Ultravioleta Internacional ha convencido a Grzegorz Chlewicki, de la Universidad de Leiden, y al autor, de que los granos interestelares se distribuyen en tres categorías diferentes. Las partículas responsables del máximo de 2200 angström en la curva de extinción deben tener un radio inferior a la millonésima de centímetro. Se cree que son partículas de carbono, semejantes quizás al grafito. Las partículas responsables de la extinción en el ultravioleta lejano deben ser también muy pequeñas; las responsables de la extinción en el dominio visible y en el infrarrojo serán, al menos, unas diez veces mayores, de un radio de una cienmilésima de centímetro, es decir, el tamaño de las partículas del humo de cigarrillo. Los únicos ele-

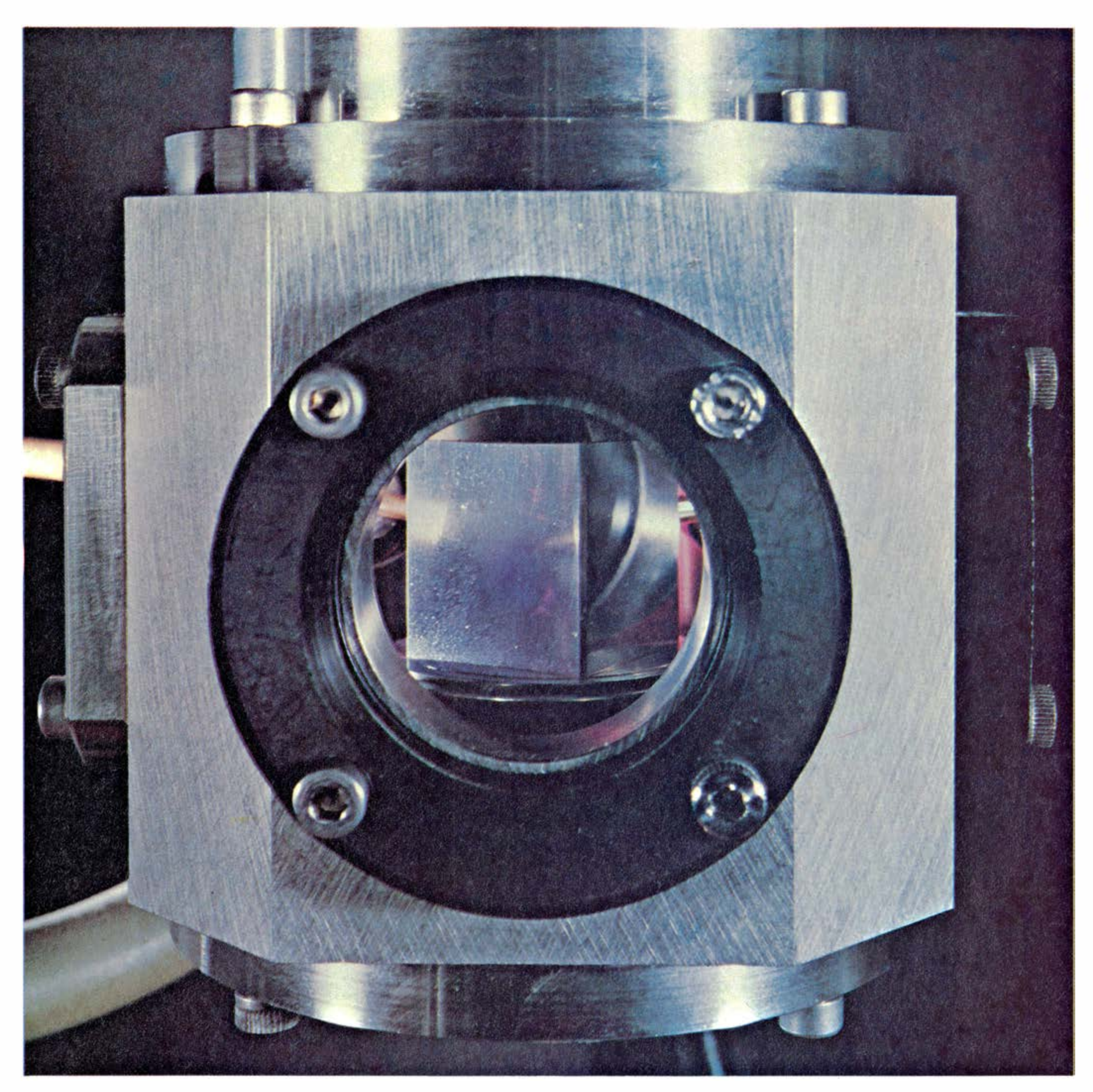


1. GRANOS INTERESTELARES, en un entramado de nubes flameantes que constan de granos y materia gaseosa, en esta fotografía de las Pléyades, cúmulo estelar situado a 400 años-luz del sistema solar, en la constelación de Tauro. Los granos hacen visibles las nubes al reflejar la luz de estrellas. La mancha brillante de la parte inferior es la nube de Merope; con mil átomos por centímetro cúbico, constituye una de las más densas de la Vía Láctea. La estrella Merope está atravesando la nube a una velocidad cercana a 30 km por segundo. La estrella cesará, pues, de alumbrar la nube, y su contenido de granos y materia gaseosa desaparecerá de la vista, dentro de un millón de años. (Fotografía tomada con el telescopio de cuatro metros de Kitt Peak.)

mentos condensables que abundan en cantidad suficiente por el espacio interestelar para constituir estas partículas mayores son los orgánicos.

Juzgadas con criterios terrestres, las partículas mayores sobresalen por su escasez. Habría que explorar en un cubo de espacio interestelar que tuviera una arista con la longitud de un campo de fútbol para hallar, en promedio, un grano de esa clase. Sin embargo, la luz de las estrellas que atraviesa una nube molecular de tamaño y densidad moderados puede fácilmente atenuarse a una cienmilésima parte a causa de la difusión (y también de la absorción) por los granos. Si el sistema solar atravesase una de tales nubes, habría una época de la Tierra en que las estrellas no serían visibles. Se vería sólo los planetas. Curiosamente, el cielo no sería oscuro, sino claro, gracias a la luz solar difundida por los granos.

La primera sospecha quizá de que los granos interestelares podían ser materia condensada en el espacio surgió, en 1935, en la mente del astrónomo sueco Bertil Lindblad. Catorce años después, el astrónomo holandés H. C. van de Hulst señaló la posibilidad de que los condensados fueran hielos. Faltaba un eslabón en el argumento: a la baja presión interestelar, no podrían condensarse hielos, excepto sobre núcleos de condensación preexistentes.



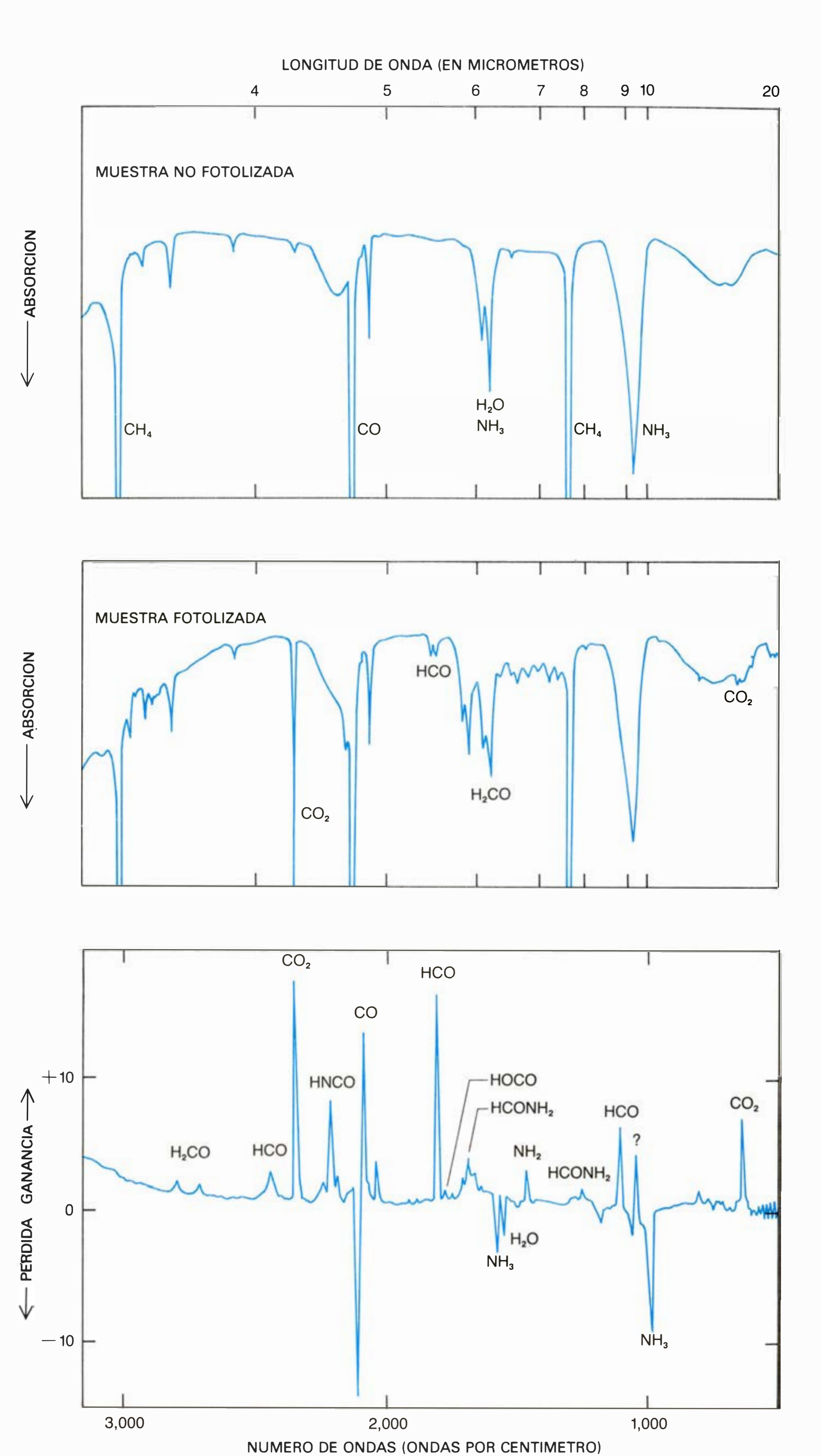
2. SIMULACION EN EL LABORATORIO de las fases iniciales del crecimiento de los granos interestelares, realizada en una cámara experimental en el laboratorio del autor, en la Universidad de Leiden. (Intervinieron también de manera destacada sus colegas Louis Allamandola y Fred Baas.) Se introdujo, en la cámara, una mezcla de agua (H_2O), metano (CH_4), amoniaco (NH_3) y otras moléculas sencillas; los elementos en esos gases representan la combinación de átomos de oxígeno, carbono y nitrógeno que se realiza en las estrellas, con átomos de hidrógeno ya presentes en el espacio interestelar. Se permitió a la mezcla depositarse sobre la superficie de un "dedo frío", enfriado hasta una temperatura de 10 grados Kelvin. Se mantenía un alto vacío. Mientras tanto, la cámara se bañaba en luz ultravioleta. El color azul de la materia en el dedo frío se debe a la creación de radicales formilo (HCO).

A principios de los años setenta, quedó establecido que las gigantes *M*, estrellas grandes y frías, y las supergigantes estaban rodeadas de pequeñas partículas de silicatos. Investigaciones posteriores sugirieron que las partículas se condensaban en la atmósfera de tales estrellas, para salir expelidas después hacia el espacio, probablemente por la presión de la radiación que emiten las estrellas. Las partículas de silicatos no tardan en alejarse de las estrellas más cercanas lo suficiente para enfriarse hasta unos 10 grados Kelvin. Ellas constituyen, pues, las semillas necesarias para el crecimiento de los granos interestelares.

Finalmente, los núcleos son capturados por una nube que tiene una densidad apreciable de átomos y moléculas gaseosas. Estas últimas se sueldan a los núcleos y se congelan, formando gradualmente una envoltura de hielos. Cabe imaginarse cuál será la composición de la envoltura. Tómense los elementos condensables más abundantes: oxígeno, carbono y nitrógeno, y supóngase que en la superficie de un núcleo de silicato se combinan con átomos de hidrógeno. El resultado es, principalmente, hielos de agua (H_2O), metano (CH_4) y amoniaco (NH_3).

El hielo de agua encierra su interés. Si abundara en el espacio interestelar, deberíamos detectarlo. En particular, el astrónomo de los años sesenta tenía derecho a esperar que la mejora en las técnicas de detección del infrarrojo haría pronto observable la absorción de radiación infrarroja a una longitud de onda de unos tres micrometros. La absorción estimula la vibración del enlace O-H en el hielo de agua. Los negativos resultados iniciales, disponibles ya en 1968, decepcionaron. A modo de explicación, podemos culpar a la fotolisis, la rotura de las moléculas de H_2O por la radiación ultravioleta incidente sobre los granos interestelares.

Por la época en que se estaba investigando, sin éxito, la presencia de hielo de agua, la detección de amoniaco y formaldehído (H_2CO) en nubes interestelares demostraba que la química del espacio era mucho más complicada de lo que se había imaginado. Químicos y astrofísicos comenzaron a buscar mecanismos que facilitasen las interacciones entre moléculas y átomos del espacio. La búsqueda ignoraba la posible relación entre la formación de moléculas complicadas y la ausencia de hielo de agua. Si las de agua podían romperse en un grano interestelar, lo mismo podía ocurrir con otras moléculas. Una sucesión de recombinaciones en el



3. ABSORCION DE RADIACION INFRARROJA por la materia del dedo frío. Revela sus cambios de composición a medida que la radiación ultravioleta modifica la mezcla. El espectro superior muestra la absorción infrarroja de una mezcla de monóxido de carbono, amoniaco, agua y metano. La mezcla no se ha irradiado; los máximos se deben solamente a las moléculas iniciales. Cada máximo representa la absorción de una longitud de onda particular de la radiación infrarroja para estimular el alargamiento, flexión u oscilación de una clase particular de moléculas. El espectro del centro muestra la absorción infrarroja de la mezcla después de la irradiación. Se han formado nuevas moléculas; entre ellas cabe destacar el dióxido de carbono, el formaldehído (H_2CO) y el radical fórmilo (HCO). Cierto número de máximos no rotulados indica otras moléculas recién formadas que todavía no se han identificado. La curva inferior representa la relación de los dos espectros de arriba. Las moléculas destruidas por la irradiación muestran mínimos, mientras que las producidas por la irradiación presentan máximos.

grano podrían crear entonces grandes moléculas a partir de las pequeñas.

Supongamos que la envoltura de un grano está formada, al principio, por una mezcla congelada de moléculas de agua, metano y amoniaco (mezcla que el autor llama hielo sucio); aceptemos, asimismo, que el grano tenga un radio de 0,12 micrometros, unos 10^{-5} centímetros. Resulta que el flujo de fotones (cuantos de radiación electromagnética) ultravioleta del espacio interestelar basta para romper cualquier enlace molecular de la envoltura, en un período de decenas o centenares de años. Las nubes interestelares perduran entre un millón y 100 millones de años, lo que implica que la envoltura no permanecerá químicamente estática.

Lo que puede ocurrir en la envoltura en virtud de la fotolisis ultravioleta resulta bastante complejo. Los fotones ultravioleta pueden disociar las moléculas de la envoltura, dejando en su lugar radicales congelados como el OH y el NH_2. Los radicales pueden recombinarse y reconstituir las moléculas originales o formar nuevas especies. O pueden, también, quedarse en su sitio, sin reaccionar. En algún momento de su vida, la envoltura debe estar formada por una mezcla congelada de moléculas y radicales de complejidad variable, mezcla que puede tener alguna relación con el hecho de que los astrónomos hayan detectado ya más de 60 clases de moléculas en el espacio interestelar.

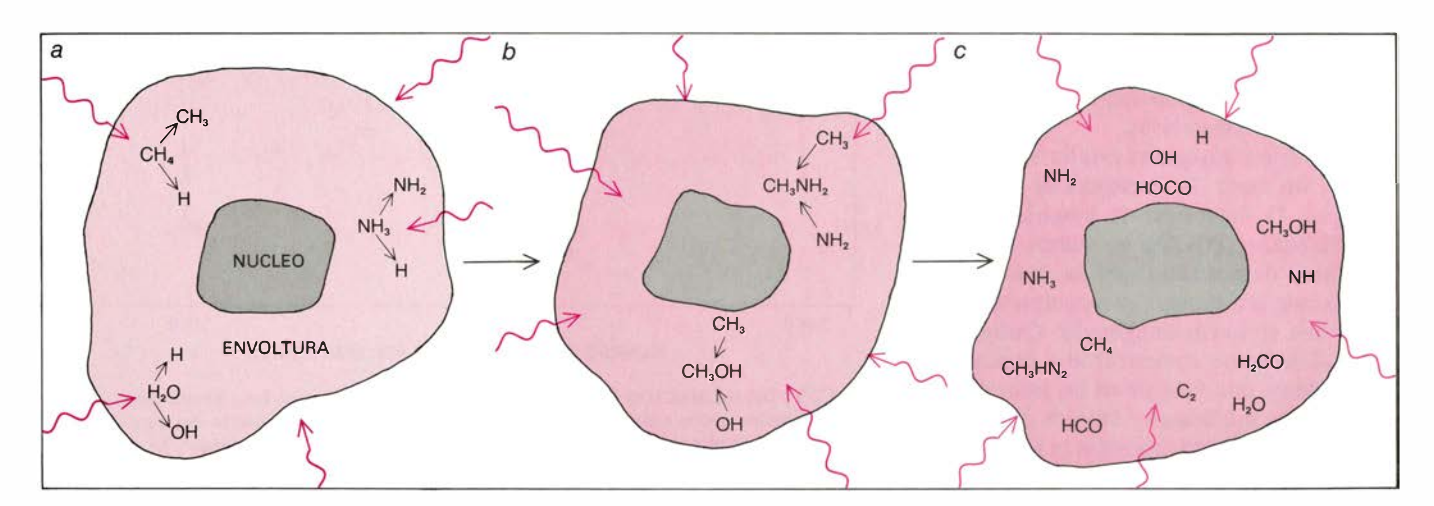
Un radical es una molécula incompleta que posee un electrón no compartido; y es, por ello, altamente reactiva. Así, cuando los radicales entran en contacto, tienden a combinar fácilmente, liberando una apreciable cantidad de energía, del orden de cuatro o cinco electronvolt. Supongamos que una milésima parte de las moléculas estables de la envoltura de un grano interestelar se formasen todas de una vez, por recombinación de radicales. La energía liberada sería suficiente para elevar la temperatura del grano en, como mínimo, 20 grados. En un grano interestelar, un suceso desencadenante tal podría conducir a una reacción en cadena en la que el calor generado por la recombinación de los radicales liberase a otros radicales congelados, permitiéndoles difundirse a través de la envoltura sólida y hallar radicales con los cuales reaccionar. De hecho, la sucesión podría producir una explosión en el grano. Unicamente el hecho de que la temperatura media de un grano sea solamente de diez grados Kelvin impide, ordinariamente, a los radicales difundirse y dar comienzo a una reacción en cadena.

La descripción anterior es una explicación simplificada de lo que debe ocurrir en un grano interestelar. Una explicación más completa requeriría conocer mejor los procesos que se desarrollan en una mezcla molecular a baja temperatura, sometida a la radiación ultravioleta. Parece, pues, necesaria una investigación de laboratorio.

Los primeros pasos en esta dirección los dieron Carl Sagan y B. N. Khare, de la Universidad de Cornell, y Andrew J. Yencha y el autor, en la Universidad estatal de Nueva York en Albany. Más tarde, el autor aceptó una plaza en Leiden, y allí, con la creación del laboratorio de astrofísica, en 1975, él y sus colegas (principalmente Louis Allamandola y Fred Baas) lograron simular en una cámara las condiciones que, en el espacio, rigen la evolución de los granos interestelares.

Los elementos clave de la simulación son la baja temperatura y la radiación ultravioleta. La baja temperatura (de hasta 10 grados Kelvin) se alcanza mediante un criostato de helio líquido que incluye un "dedo frío", que puede ser un bloque de aluminio o una ventana transparente montada en un anillo metálico. Por una ventana se introduce en la cámara la radiación ultravioleta. Gases varios penetran a través de un tubo. La presión en la cámara puede hacerse rebajar hasta 10^{-8} torr. (La presión en el espacio interestelar es de 10^{-19} torr, y menor incluso.) Por regla general, los gases son mezclas de moléculas volátiles sencillas: CH_4, CO, H_2O, CO_2, NH_3, N_2 y O_2. Los gases se condensan sobre el dedo frío, que actúa, por consiguiente, como el núcleo de un grano interestelar. Mientras tanto, se irradia el condensado.

La diferencia más importante entre la cámara y el espacio interestelar es la escala de tiempo de la fotolisis. Una hora de irradiación en el laboratorio equivale a 1000 años de irradiación en una nube interestelar difusa. Durante el experimento, a través de una segunda ventana, podemos dirigir el haz de un espectrofotómetro infrarrojo que mide la absorción infrarroja de la muestra, a longitudes de onda entre 2,5 y 25 micrometros. En esta "región de identificación", se identifican las moléculas por sus modos de oscilación de alargamiento, curvatura y bamboleo en un sólido. Otros sensores miden la presión, luminiscencia, las masas de las moléculas gaseosas y la absorción de luz visible.



4. ESTRUCTURA DE LOS GRANOS cuando se depositan inicialmente. Se infiere de las simulaciones de laboratorio. En éstas, mezclas de agua, metano, amoniaco y otras moléculas sencillas se someten a irradiación ultravioleta, a temperaturas de 10 grados Kelvin. Cada grano comienza como un núcleo de silicato que se condensó en la atmósfera de una estrella gigante fría. Alrededor de este núcleo se forma una envoltura de hielos. La radiación ultravioleta rompe algunas moléculas de la envoltura, formando radicales, o fragmentos moleculares reactivos (*a*). Los radicales pueden recombinarse de nuevas maneras (*b*). En un período de tiempo más largo, la irradiación ultravioleta continuada del grano origina mezclas de moléculas y radicales más complejos (*c*).

La creación de nuevas moléculas y radicales en cualquier mezcla inicial se detecta fácilmente a partir de las variaciones en la absorción infrarroja a medida que el experimento prosigue. Así, si comenzamos con una mezcla de CO, NH_3, H_2O y CH_4, encontramos que, casi inmediatamente, aparece el CO_2, junto con formaldehído y una apreciable proporción del radical formilo (HCO). Hay otros muchos máximos en el espectro de absorción sin identificar, aunque, con toda seguridad, manifiestan la presencia de más moléculas recién formadas.

En un campo en rápido desarrollo, la observación y la teoría tienden a progresar a saltos. No mucho después del fracaso en la detección de hielo de agua en el espacio interestelar, y poco antes de la simulación con éxito de las condiciones interestelares en el laboratorio, se observó una intensa absorción astronómica a una longitud de onda de 3,1 micrometros en el espectro infrarrojo de lo que se ha dado en llamar el objeto de Becklin-Neugebauer, una fuente de radiación infrarroja en la constelación de Orión. Constituía una prueba convincente de la existencia de H_2O en granos, al menos en algunas partes del espacio interestelar. El espectro ofrecía también una ancha banda de absorción a una longitud de onda de 9,7 micrometros. La atribuimos al núcleo de silicato de los granos. Hoy se supone que el objeto de Becklin-Neugebauer es un centro de formación de estrellas. El hielo de agua se ha identificado ya no sólo en esos sitios, sino también en nubes moleculares.

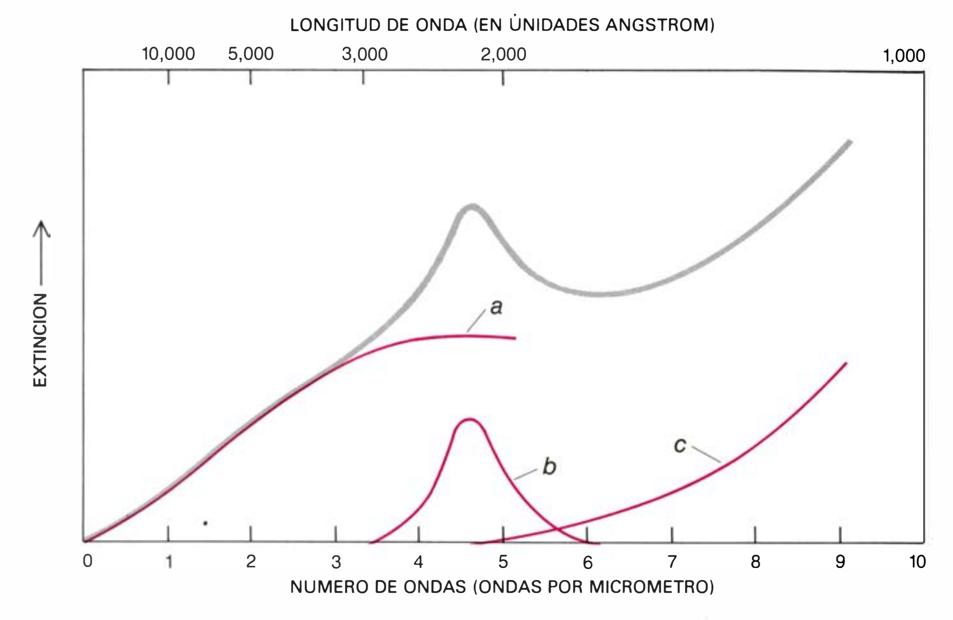
Las primeras medidas de absorción astronómica por hielo ordinario mostraban algunas discrepancias cuando se comparaban con las medidas de laboratorio de hielo de agua puro y cristalino. La absorción estaba corrida, en cuanto a la longitud de onda, con respecto a la absorción en el laboratorio; era, además, mucho más ancha. ¿Posee alguna peculiaridad la estructura cristalina del agua en los granos interestelares? En Leiden, Allamandola, Wim Hagen y el autor decidieron investigarlo. Vimos inmediatamente que el agua pura, depositada lentamente sobre un dedo frío a 10 grados Kelvin, mostraba una absorción infrarroja que presentaba una sorprendente semejanza con la absorción astronómica a tres micrometros. Se hallaba en la posición correcta del espectro, con la anchura que le correspondía. También vimos por qué. El hielo que se forma a temperaturas de 10 grados es altamente amorfo: su estructura no repite la distribución cristalina regular del hielo ordinario; antes bien, sugiere que las moléculas del líquido, moviéndose al azar, se han congelado en el camino. A medida que sube la temperatura, el hielo se torna menos amorfo, hasta que, por encima de los 130 grados, su estructura cristalina adquiere plena regularidad. Basta incluso una temperatura de 80 grados para hacer que la absorción a tres micrometros se vuelva apreciablemente más estrecha y profunda de lo que era.

También hallamos que una mezcla de agua con algo de metano y amoniaco producía una extensión, o "ala", hacia el sector de onda larga en la curva de absorción. Logramos así simular un aspecto más de la absorción astronómica. Las comparaciones entre los espectros astronómicos y los espectros de los hielos puros e impuros, a diversas temperaturas, nos proporcionaron los datos que necesitábamos para seguir la evolución de los mantos de los gránulos en el espacio interestelar.

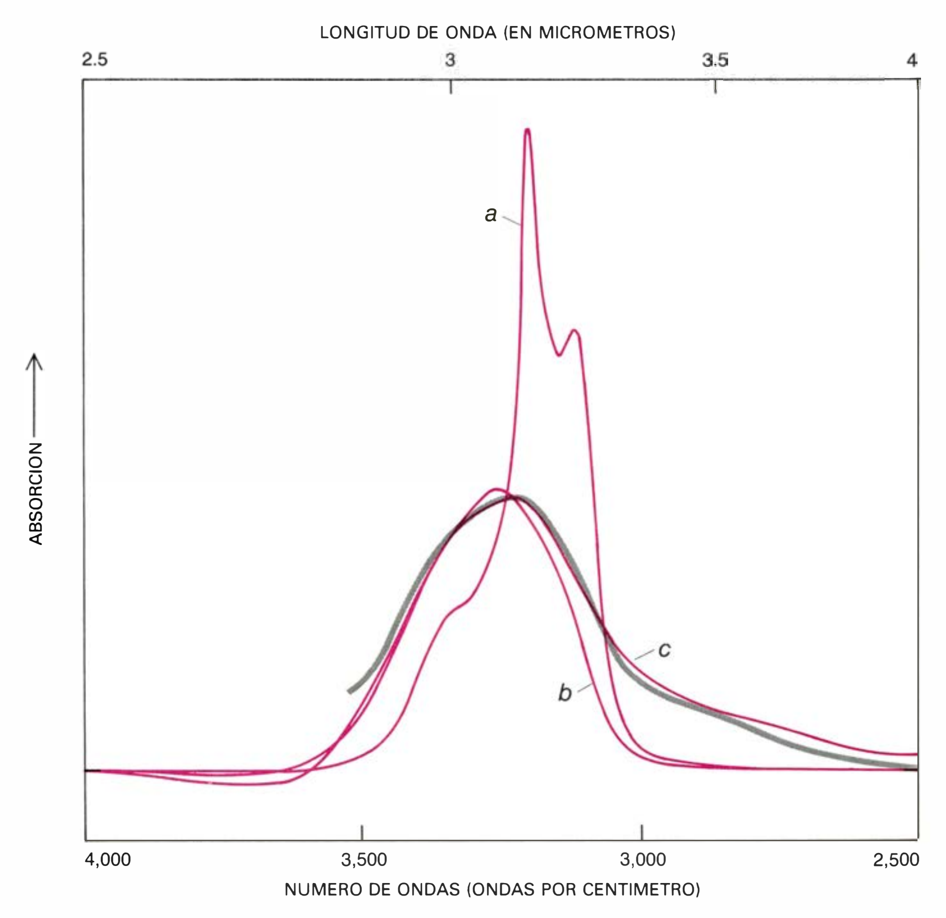
¿De qué están compuestos los granos interestelares cuando no contienen agua? Para averiguarlo, el autor y sus colegas siguieron la evolución de granos más allá de las fases en que se formaban por deposición, se fotolizaban por radiación ultravioleta y mantenían a temperaturas extremadamente bajas en el espacio interestelar. El período de inactividad a baja temperatura llena, sin duda, la mayor parte de la vida de un grano. Pese a ello, debe haber muchas situaciones que conducen al calentamiento del grano. Un ejemplo obvio es la formación de nuevas estrellas en una región del espacio rica en granos interestelares.

Acometimos la estrategia siguiente. Durante la deposición de los gases sobre el dedo frío y la irradiación ultravioleta simultánea de la creciente capa de envoltura, mantenemos una temperatura baja y constante, entre 10 y 15 grados. La capa se hace más espesa. Para velocidades de deposición suficientemente bajas, la capa puede ser amorfa, esto es, vidriosa; adquiere un brillante color azul, debido a la creciente concentración de radicales fórmilo. Si se detiene la deposición, desconectando la lámpara ultravioleta y manteniendo una temperatura constante, no se observan cambios ulteriores. El manto permanece completamente estable durante períodos de hasta una semana.

Si dejamos que la temperatura ascienda lentamente, la envoltura emite una luz verde azulada. Trabajando en el laboratorio del autor, Leo van Ijzendoorn ha atribuido la luz a la liberación de energía por pares de radicales formilo al combinarse para formar moléculas de glioxal, es decir, dos radicales HCO unidos por sus moléculas de carbono. Si hay monóxido de carbono en la en-



5. ATENUACION DE LA LUZ DE LAS ESTRELLAS por los granos interestelares, resultado de la difusión selectiva de ciertas longitudes de onda. (Una partícula sólida difunde más eficazmente a longitudes de onda parecidas a su propio diámetro.) El efecto de conjunto es hacer que las estrellas parezcan algo más rojas de lo que son. Se cree que la curva de extinción desde el infrarrojo lejano (*izquierda*) al ultravioleta extremo (*derecha*) tiene tres componentes. El crecimiento a la izquierda (*a*) se atribuye a partículas de un radio de 10^{-5} centímetros. El máximo central, a 2200 unidades angström, (*b*), se atribuye a partículas diez veces menores, formadas de carbono puro, muy parecido seguramente al grafito. El crecimiento que se registra a la derecha (*c*) se atribuye a partículas igualmente pequeñas, constituidas, quizá, por silicatos.



6. ESPECTRO DEL OBJETO DE BECKLIN-NEUGEBAUER, fuente infrarroja en la constelación de Orión, que marcó la primera detección de agua en el espacio interestelar. El espectro se muestra en gris. Presenta un máximo a una longitud de onda en el infrarrojo de unos 3,1 micrometros. La absorción de luz a esa longitud de onda estimula la vibración de un enlace O-H en el hielo de agua. Se ofrece, para comparación, tres espectros de laboratorio. El del hielo de agua cristalino (*a*) se ajusta bastante mal a la curva astronómica: su máximo de absorción a 3,1 micrometros resulta demasiado abrupto y estrecho. La absorción está además corrida en longitud de onda. El espectro del hielo de agua amorfo que se depositó sobre el dedo frío a 10 grados Kelvin (*b*) se ajusta bastante mejor a la curva. El espectro de agua, amoniaco y metano en las relaciones de seis a uno y a tres (*c*) es el que mejor se ajusta de todos. El objeto de Becklin-Neugebauer puede ser una densa nube interestelar en la que se estén formando estrellas.

voltura, las moléculas de CO se combinan con átomos de hidrógeno aumentando la concentración de HCO. El próximo paso en el proceso es la formación de formaldehído (H_2CO), además de glioxal.

La luz verde azulada representa sólo una pequeña parte de la energía liberada cuando se combinan los radicales. El resto se invierte en calentar la muestra. Con un calentamiento rápido, hasta el punto de que un gran número de radicales queden desprendidos de sus lugares, puede explotar la muestra. Algunos aspectos de esas explosiones han sido revelados mediante films de varias muestras, realizados por Louis d'Hendecourt. La explosión principal se produce en una parte de la envoltura donde la temperatura ha subido hasta unos 25 grados. Allí, la materia de la envoltura sale proyectada desde el dedo frío, dejando expuesta la superficie. En otros lugares, la materia continúa adherida y se siguen viendo pequeños relámpagos, tanto locales como difusos, demostrando que los radicales pueden sobrevivir hasta que la muestra alcanza una temperatura de unos 100 grados. Asistimos a una observación sorprendente: los fotones ultravioleta crean radicales en la envoltura con un rendimiento que alcanza el 10 por ciento, al menos hasta que dicho manto comienza a saturarse con radicales. Explicado de otra manera: bastan 10 fotones para producir un incremento neto de un radical en la mezcla.

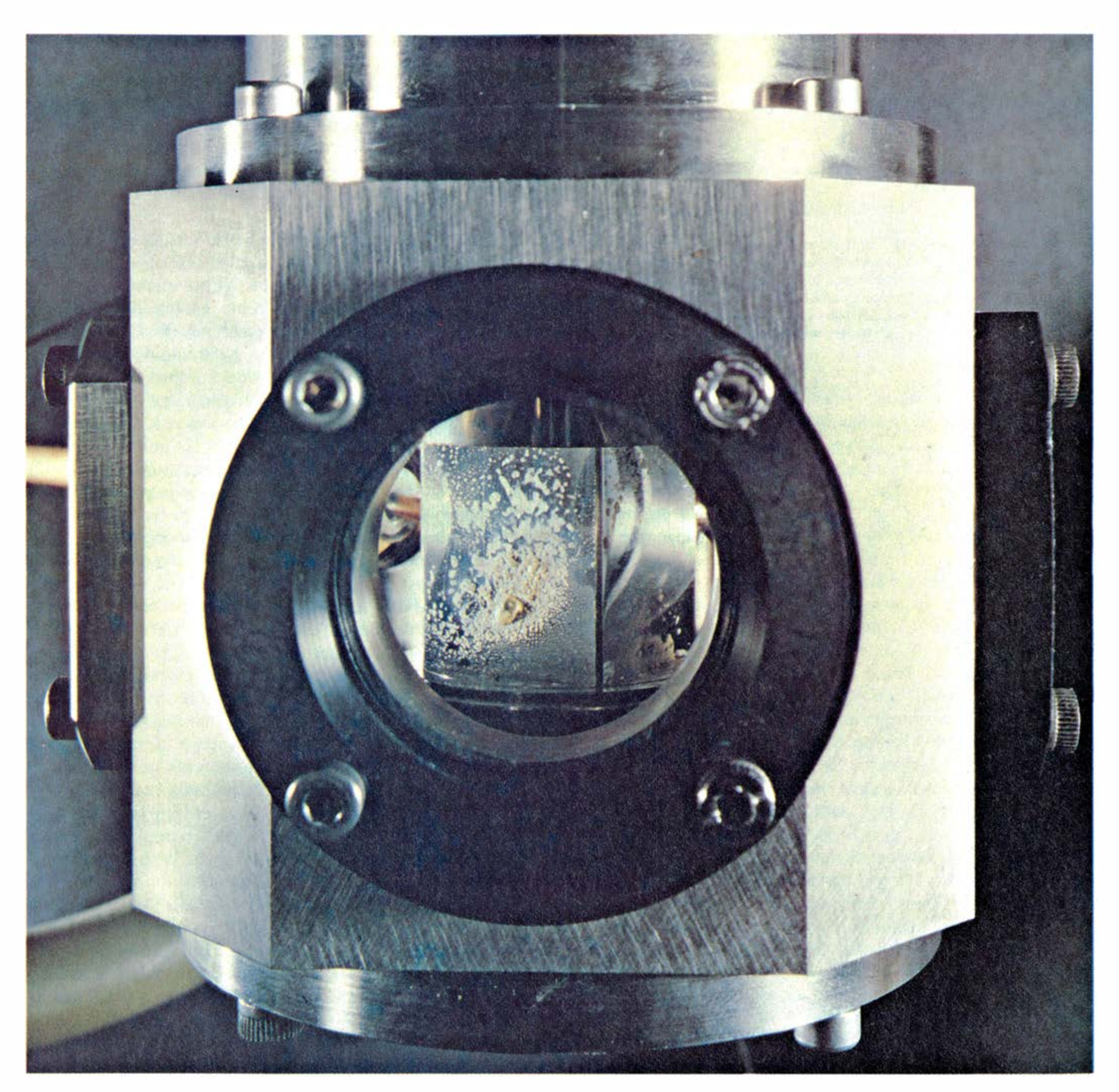
Cuando una muestra bien irradiada y bien fotolizada se calienta con lentitud suficiente para evitar reacciones incrontroladas y las explosiones que producen, queda un substancial residuo amarillento que sobrevive indefinidamente, incluso a temperatura ambiente (295 grados Kelvin). Se cree que el residuo es semejante a la mezcla de complejas moléculas orgánicas sintetizada por la fotolisis de los gránulos en el espacio interestelar. De hecho, este trabajo se parece, en cierto modo, al de Stanley L. Miller y Harold C. Urey, quienes simularon la producción de moléculas orgánicas en la atmósfera de la primitiva Tierra, aplicando descargas eléctricas a mezclas gaseosas de agua, metano, amoniaco y otras moléculas.

La diferencia entre el flujo de fotones ultravioleta en el laboratorio y el flujo en el espacio interestelar conduce a una estimación de la velocidad a la cual la fotolisis en el espacio convierte moléculas volátiles simples en moléculas orgánicas complicadas. Entre un uno y un 10 por ciento de las moléculas sencillas de una nube interestelar se transforma en moléculas complejas en 10 millones de años. He aquí, pues, la explicación de los granos interestelares que carecen de hielo de agua. El hielo, junto con otras moléculas pequeñas, se ha convertido en substancias más complejas.

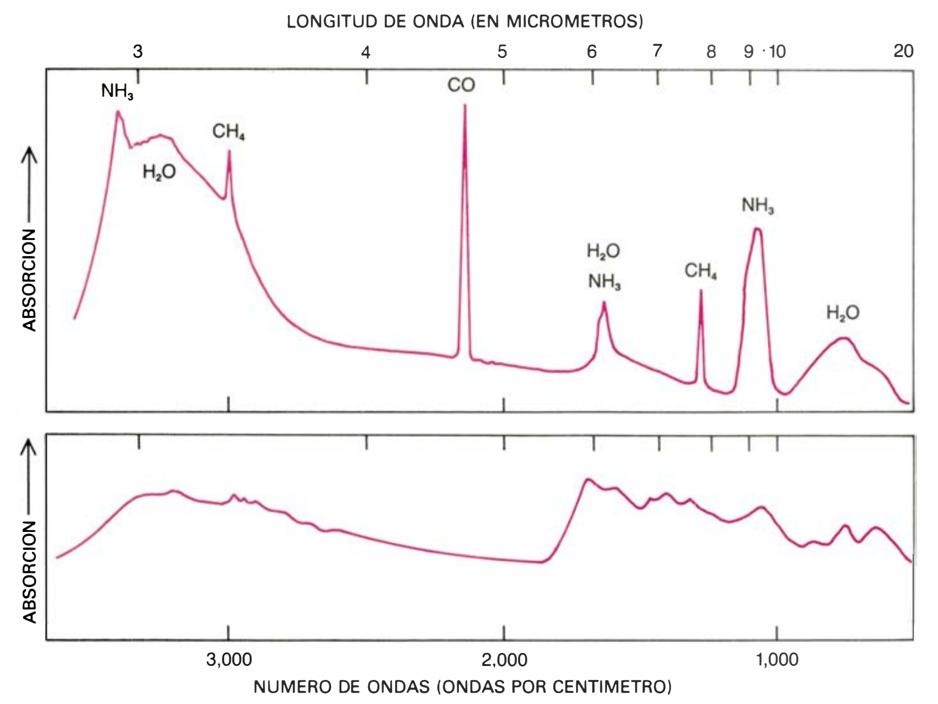
El análisis e identificación química del residuo amarillo creado en el laboratorio –le llamamos materia amarilla– ha tropezado con la dificultad de la cantidad obtenida en cada experimento, pequeña. Sin embargo, se ha observado que, juzgado con criterio astronómico, el residuo es muy poco volátil y, consiguientemente, puede sobrevivir a condiciones mucho más duras de las que podría soportar un manto de hielo. Por ejemplo, el hielo de agua se evapora en el espacio a unos 90 grados Kelvin, mientras que la materia amarilla no lo hace a temperaturas inferiores a 450 grados. Además, tanto los análisis físicos como los químicos revelan que el residuo está notablemente enriquecido en carbono. La espectroscopía de masas de un residuo, realizada por H. J. de Jong, en Leiden, mostró que la abundancia del carbono duplicaba la del oxígeno. En el medio interestelar (y en nuestra mezcla inicial de gases), el oxígeno es dos veces más abundante que el carbono; por tanto, en la materia amarilla, la fracción de oxígeno se ha reducido a la cuarta parte. Se está desarrollando, en Leiden, un estudio sistemático de la química de la materia amarilla, en colaboración con James P. Ferris, del Instituto Politécnico Rensselaer. Mientras tanto, se está creando, en Leiden, residuos de complicación creciente. Se ha iniciado, también, la irradiación ultravioleta a largo plazo, a fin de seguir su evolución ulterior bajo condiciones interestelares simuladas.

Apoyados en los resultados de nuestro laboratorio, nos hallamos ahora en situación de colocar las observaciones astronómicas en un marco teórico y proponer un esquema para la evolución, tanto de los gránulos interestelares como de las nubes gaseosas en las que se encuentran. Considérese primero las nubes. Su evolución a gran escala depende de sucesos que tienen lugar en la galaxia. Las nubes colisionan entre sí e interaccionan con estrellas; a escala mayor, quedan modificadas por las fuerzas hidrodinámicas y gravitatorias inducidas al girar en torno al eje de la galaxia. Presumiblemente, las nubes densas se forman a partir de condensación de las difusas. Después, las estrellas se constituyen en las nubes más densas, por contracción gravitatoria. Cuando las estrellas comienzan a generar energía por fusión termonuclear, la materia restante de la nube vuelve al espacio interestelar. Frecuentemente se proyecta por ondas de choque y vientos de gas y partículas, emitidos por las estrellas recién nacidas. Calentada y sometida a baja presión, la materia restante se expande. El resultado es una nube interestelar difusa. Se cierra así un ciclo.

El ciclo se repite una y otra vez. Mas, siempre que se forma una estrella, debe consumirse una fracción del polvo y gas interestelares. Los cálculos actuales cifran el nacimiento de una a dos estrellas por año en la Vía Láctea; de aquí puede deducirse que la totalidad del medio interestelar se consumiría en unos 5×10^9 años, aproximadamente la edad de la Tierra. En promedio, pues, ningún grano interestelar sobrevivirá



7. "MATERIA AMARILLA" producida por la lenta irradiación de la helada envoltura de los granos interestelares; aquí aparece sobre la superficie de un dedo frío, en el laboratorio del autor. El análisis de la materia se hace difícil, debido a la pequeña cantidad que se produce en cada experimento. En general, la materia amarilla constituye una mezcla complicada de moléculas orgánicas notablemente enriquecidas en carbono, con respecto a la composición media del espacio interestelar. Muéstrase notablemente estable, incluso a la temperatura ambiente. Puede sobrevivir a las duras condiciones espaciales.



8. ESPECTRO INFRARROJO de la materia amarilla (*arriba*), comparado con un espectro de una mezcla que contenga las mismas proporciones de oxígeno, carbono y nitrógeno (*abajo*). Este último espectro presenta una prominente absorción por hielo de agua a 3,1 micrometros; el primero carece de tal máximo. El oxígeno de la materia amarilla está ligado en las moléculas orgánicas hasta tal punto que no queda nada para combinarse con el hidrógeno, produciendo agua. La abundancia de materia amarilla en los granos interestelares maduros puede explicar bien por qué razón dichos granos no presentan agua.

un tiempo mayor. Sin embargo, cada fase de la evolución de una nube no dura más allá de unos 10^8 años. Así, un grano puede tener una rica y variada historia, pasando por dentro y fuera de nubes densas 25 o más veces, hasta incorporarse finalmente a una estrella.

Sigamos la historia de un grano a través de uno de sus ciclos. Comienza en una nube difusa. Cabe suponer que es un grano que ya ha experimentado, al menos, un ciclo desde su primera aparición como núcleo de silicato en la atmósfera de una antigua estrella supergigante. Así, el grano tiene un manto, y como su espectro de absorción infrarrojo no revela agua, suponemos que la envoltura está constituida por materia amarilla. El cálculo muestra que esos gránulos con núcleo y envoltura (junto con las partículas más pequeñas de carbono y silicatos, responsables de la atenuación de la luz de las estrellas en las longitudes de onda del ultravioleta) dan cuenta de, aproximadamente, el 20 por ciento de todo el oxígeno de la galaxia, junto con el 70 por ciento del carbono y cerca del 100 por ciento del silicio, magnesio e hierro. Lo que queda en forma gaseosa en la nube difiere, pues, de una mezcla de elementos químicos en sus abundancias galácticas relativas. Las observaciones de las nubes difusas confirman esta hipótesis. La relación de oxígeno a carbono en una nube difusa es de cinco a uno.

La abundancia de oxígeno gaseoso sugiere que, cuando la nube se ha hecho lo suficientemente densa para permitir que se formen moléculas, el agua gaseosa debería ser abundante. ¿Por qué no lo es? ¿Por qué no se aprecia agua? Las observaciones de granos de polvo en longitudes de onda del infrarrojo realizadas por Douglas Whittet y sus colegas en el Politécnico de Preston, en Gran Bretaña, muestran que los granos interestelares en las nubes moleculares ordinarias (no necesariamente en las que se están formando estrellas) muestran señales de absorción por hielo de agua a tres micrometros. Evidentemente, se está formando agua en los gránulos con núcleo y envoltura. Estamos estudiando esta posibilidad en Leiden. En particular, Xander Tielens y d'Hendecourt han formulado un esquema de reacciones químicas que abarca no sólo las reacciones de la fase gaseosa, sino también de la superficie e interior de los granos. Ello nos lleva a describir un grano en una nube molecular como algo que tiene típicamente dos envolturas: un manto interior de substancia amarilla y otro exterior que encierra una proporción destacada de agua.

El aumento de tamaño de un grano interestelar debido al crecimiento de un manto externo debe alterar la manera como extingue la luz visible. Sería de esperar, pues, una correlación entre la pauta de extinción visual y la intensidad de la banda de absorción de hielo a tres micrometros. Las observaciones astronómicas lo confirman en cierto número de casos. Además, los cálculos teóricos realizados, en Leiden, por C. E. P. M. van de Bult y el autor muestran que los granos de polvo en las nubes protoestelares, tales como el objeto de Becklin-Neugebauer, y los granos de nubes más tenues poseen hasta un 60 por ciento de su envoltura exterior en forma de agua. Más aún, los estudios de Gerard van de Zwet y Baas proporcionan un esperanzador apoyo a la presencia de moléculas de H_2CO en la fase sólida. Teniendo en cuenta todo ello, la existencia de la helada envoltura exterior parece bastante bien establecida.

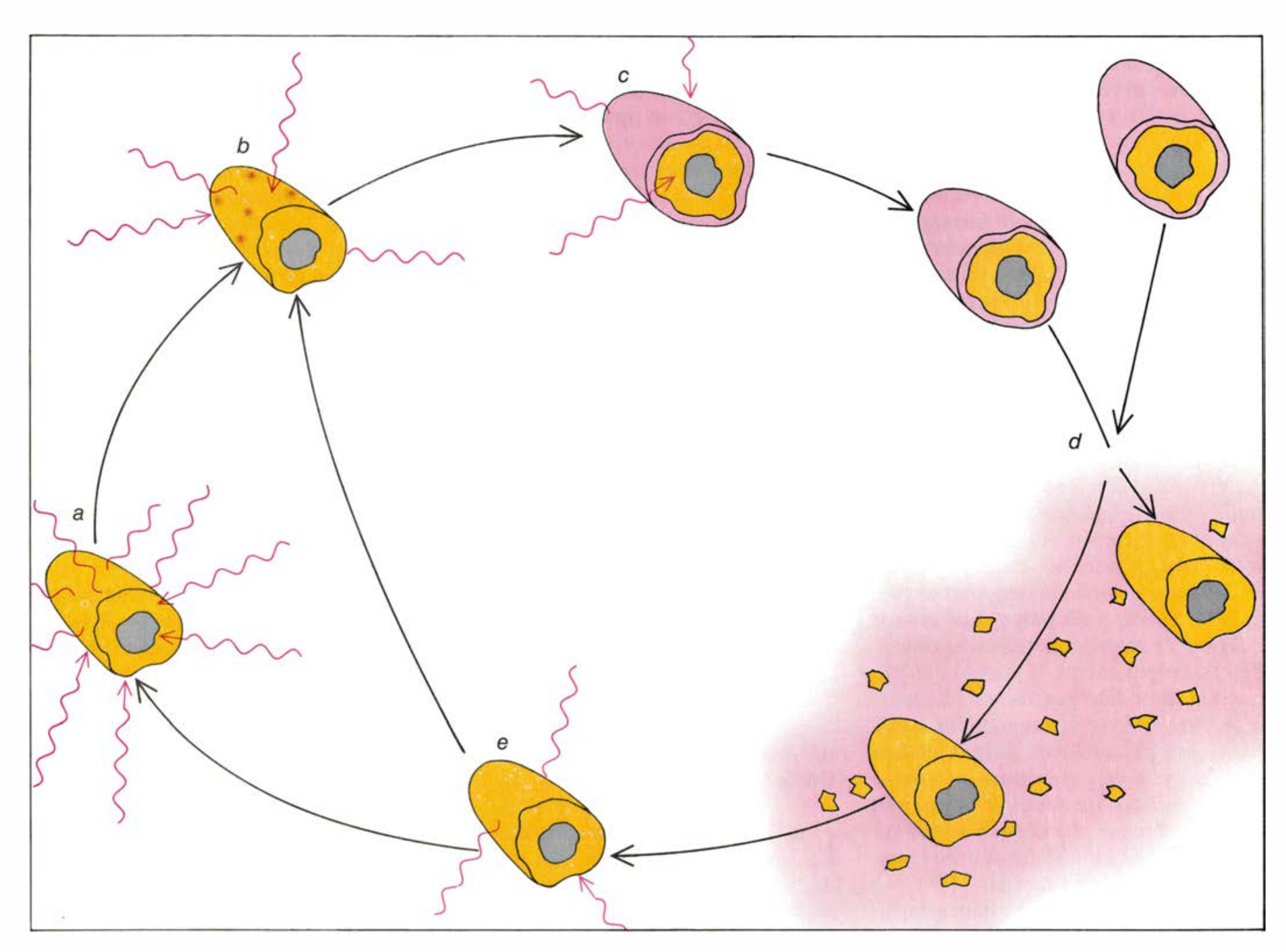
La existencia de una envoltura exterior conduce, por otra parte, a un rompecabezas. ¿Qué impide que la envoltura exterior crezca hasta agotar todo el gas condensable de una nube? En particular, ¿qué impide que las moléculas gaseosas se adhieran a la envoltura? O si lo hacen, ¿qué es lo que las arranca? Se han avanzado muchas sugerencias, pero en las nubes moleculares sólo hay un mecanismo de eficacia suficiente para cobrar importancia. Me refiero a la explosión de la envoltura del grano. El proceso comienza cuando, al principio, el grano desarrolla una envoltura por captura de átomos y moléculas. Durante el proceso de captación, el grano está sometido a un bombardeo de fotones ultravioleta. Así, la envoltura acumula cierta cantidad de radicales almacenados en tanto permanezca extremadamente frío. Ocurre, entonces, un suceso desencadenante que eleva bruscamente la temperatura del grano. De acuerdo con las simulaciones realizadas en nuestro laboratorio, basta un aumento de sólo 15 grados Kelvin para producir la detonación.

Una colisión de dos granos a una velocidad relativa de 50 a 100 metros por segundo resulta un excelente agente desencadenante. Tal velocidad es mil veces mayor que la velocidad "térmica" característica de un cuerpo sólido en una nube a 100 grados Kelvin, por ejemplo. Sin embargo, la velocidad es posible: la genera una turbulencia en la nube. (El ensanchamiento de las bandas de absorción en las nubes moleculares confirma que el gas de una nube forma remolinos a velocidades de miles de metros por segundo. Sin duda,

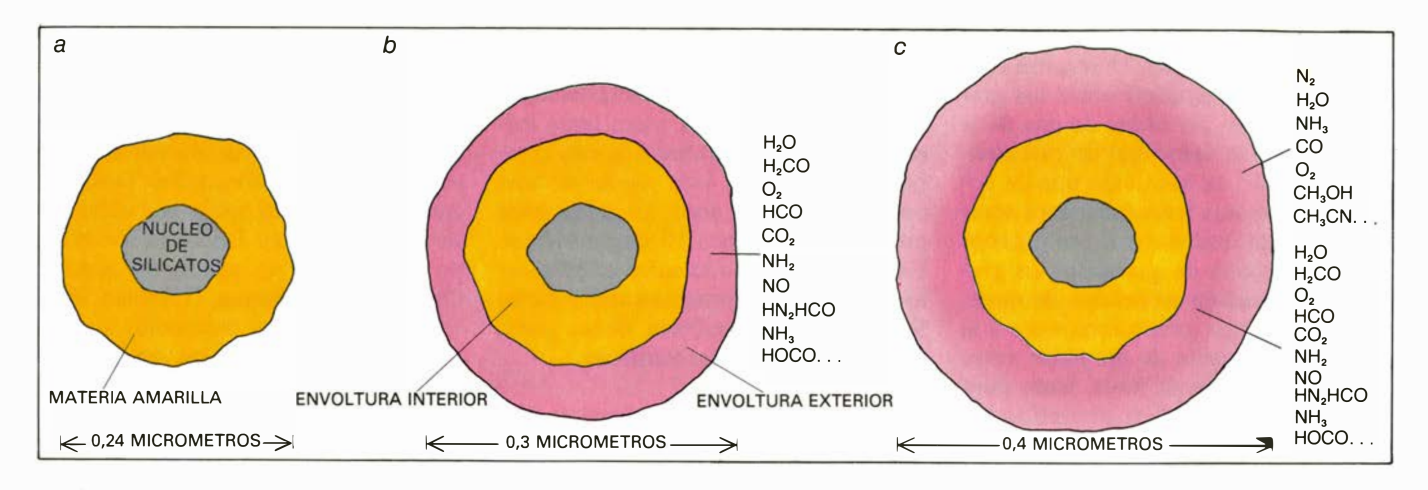
los granos son arrastrados junto con el gas; en cualquier caso, el arrastre será suficiente para producir entre los granos velocidades del orden de una décima parte de la velocidad de turbulencia.) Además, no hace falta que las colisiones sean muy frecuentes para mantener un equilibrio entre la lenta y continua deposición de gas sobre los granos y la liberación esporádica de materia del grano que pasa nuevamente a la nube. La explosión de un grano entre diez, en un millón de años, basta para mantener la densidad de moléculas observada en el estado gaseoso.

Durante su estancia en una nube molecular, el grano hace más que servir de intermediario en la formación de moléculas gaseosas. Su envoltura acumula un residuo complejo. Después de todo, por cada colisión u otro mecanismo detonante que produzca la explosión de un grano, hay muchos sucesos menos violentos, tales como una colisión suave, que originan una sucesión autolimitada de reacciones y una lenta formación de un manto todavía más complejo. Durante la vida media de una nube, de unos 10^7 años, un grano sufre probablemente unas 10 detonaciones. En el mismo período sufre quizá cientos de calentamientos no catastróficos. Su composición química debe, pues, modificarse considerablemente.

Atendamos ahora a los sucesos que tienen lugar cuando las condensaciones locales de una nube molecular, o nudos de alta densidad, proporcionan las condiciones necesarias para la contracción final de la materia en una estrella o grupo de ellas. En ese momento, la velocidad de deposición de las moléculas gaseosas sobre los granos interestelares aumenta hasta que sólo quede hidrógeno y helio por capturar. Pronto, algunos de los granos que han crecido quedan barridos por el gas que se contrae. Se convierten entonces en parte de una protoestrella. Después, cuando la estrella inicia la fusión nuclear, se evaporan. De hecho, la materia del interior se separa en átomos. Otros granos escapan. También éstos son mayores que el promedio. Sin embargo, quedan sometidos, en breve, a una gran variedad de procesos destructivos y erosivos. Por ejemplo, todos los cuerpos volátiles de un grano que se está calentando con su proximidad a una protoestrella se evaporan. Mientras, otros granos, más distantes de las protoestrellas, pueden conservar su envoltura engrosada, incluidas materias volátiles tales como H_2O. Esto explicaría la observación de que los granos de-



9. CICLO DE VIDA de los granos interestelares. Los lleva de una nube difusa a una densa y, después, si no entran a formar parte de una estrella recién nacida, otra vez a una nube difusa. En una nube difusa, los granos comienzan siendo partículas de silicatos. Aquí se muestran con una envoltura de materia amarilla procedente de un ciclo anterior (*a*). La nube se hace más densa; sobre los granos se deposita una capa de hielos a velocidad creciente (*b, c*). Con el tiempo, la nube se hace suficientemente densa para que su materia gaseosa adquiera forma molecular. La radiación ultravioleta da lugar a radicales en el helado manto de los granos; colisiones ocasionales entre los granos calentarán la envoltura para que los radicales se recombinen en gran número, liberando energía. En algunos casos, la envoltura explota, pasando a enriquecer de nuevo la fase gaseosa de la nube (*d*). En las partes más densas de las nubes, se forman estrellas con el colapso gravitatorio de polvo y gas. El calor de las estrellas recién nacidas evapora la helada envoltura de los granos. Lo que queda es materia amarilla (*e*). El grano puede repetir el ciclo dentro de una nube molecular. En otro caso, la nube puede dispersarse; o bien las nuevas estrellas la expelen de su vecindad, o bien un complejo entero de nubes y estrellas sale de un brazo espiral de la galaxia. Queda, de esta manera, cerrado el ciclo. La llegada de un cuanto ultravioleta continúa la lenta y fría reconstrucción química de la materia amarilla en el manto del grano.



10. LOS GRANOS DE NUBES DIFERENTES parecen tener distintas capas, según la fase de su ciclo de vida. Los granos de una nube difusa tienen sólo un núcleo de silicato y un manto de materia amarilla (*a*). Los granos de una nube molecular encierran un núcleo de silicato y poseen un manto de materia amarilla y un manto adicional formado de condensados congelados y modificados por la radiación ultravioleta (*b*). Los granos, en una nube tan densa que estén a punto de formarse estrellas, muestran una composición más complicada todavía (*c*). Una rápida deposición de materia les ha dado una envoltura más externa de hielos no modificados por la irradiación. Simplemente, no ha habido tiempo. La polarización de la luz que atraviesa las nubes interestelares sugiere que los granos interestelares con mantos son alargados. Cuando tales partículas están alineadas, actúan débilmente como filtros de polarización.

tectados en la dirección de una intensa fuente infrarroja (presumiblemente una protoestrella o una estrella muy joven) tienen la probabilidad máxima de mostrar una intensa banda de absorción de hielo a tres micrometros.

Un complejo de nubes moleculares y las estrellas que continúan formándose en su interior sobrevive durante unos 10^8 años. Al final de este período, el complejo emerge de un brazo espiral. (Se cree que la configuración de brazos espirales de la Vía Láctea gira como un tren de ondas de densidad a la mitad de la velocidad de rotación de la materia que compone la galaxia. De aquí que, a intervalos periódicos, una porción dada de materia ayude a formar un brazo espiral. Unos 10^8 años más tarde, adelanta al borde anterior del brazo.) La materia está, entonces, en una región entre brazos, donde la densidad es más baja, y tiende a disiparse. Los granos quedan, de nuevo, en una nube difusa. Pierden rápidamente la materia volátil que hayan podido guardar en su superficie. Además, se ven sometidos a distintos procesos erosivos. Sufren erosión cuando las nubes colisionan o cuando las explosiones de supernova generan ondas de choque. La erosión puede reducir gradualmente el espesor de la envoltura de materia amarilla.

Algunos investigadores han avanzado cálculos de la velocidad de erosión; pero, en el mejor de los casos, la física del proceso es mal conocida. El trabajo de Erwin E. Salpeter y Bruce T. Draine, en Cornell, permite suponer que la materia amarilla se destruye en las nubes difusas a la misma velocidad con que se crea en las nubes moleculares. Y podemos admitir que un grano interestelar pasa aproximadamente la mitad de su vida en nubes difusas y la otra mitad en nubes moleculares. Se concluye que un grano típico llega a alcanzar una composición media estable. Así, no debe sorprendernos que la extinción de la luz de las estrellas sea prácticamente la misma en cualquier dirección en el medio interestelar difuso.

La materia amarilla constituye un componente principal de los granos interestelares; se cifra en el 1 por ciento de toda la masa de la Vía Láctea. Resulta, pues, razonable afirmar que su masa excede la de todos los planetas de la galaxia. Un motivo más para deplorar que la materia amarilla no puedan recogerla los vehículos espaciales. Un satélite artificial que viajase a una velocidad de un kilómetro por segundo y llevase desplegada una "vela" de 100 metros cuadrados de superficie recogería sólo 10^{-11} gramos de substancia por año en nuestra proximidad interestelar. Quizá la muestra que buscamos se capture cuando un vehículo espacial recoja un trozo de cometa y lo devuelva a la Tierra.

Los cometas están constituidos por granos interestelares. En particular, los estudios teóricos de Ludwig Biermann, del Instituto Max Planck de Física y Astrofísica, en Garching, y de L. Mestel, de la Universidad de Sussex, apoyan la hipótesis según la cual los cometas se formaron casi al mismo tiempo que los planetas, si bien se agregaron directamente a partir de granos de polvo de una densa nube interestelar, ligada gravitatoriamente a la nebulosa presolar (nube a partir de la cual se condensó el sistema solar). Los cometas ocupan hoy la nube de Oort, una región que contiene de 10^{11} a 10^{12} unidades que describen órbitas al azar en torno al Sol, a distancias iguales, como máximo, a la mitad de la distancia que media con respecto a las estrellas más próximas. Los cometas que se observan cerca del Sol han sufrido perturbaciones en sus órbitas originales más lejanas. Hay un dato fundamental: los cometas nacieron fríos y fríos se han mantenido durante los 4500 millones de años transcurridos desde su nacimiento. El núcleo de un cometa es un recipiente que contiene una muestra congelada de los granos presolares interestelares. Hasta los propios cometas cuya órbita les lleva cerca del calor del Sol deben haber conservado, en gran medida, su composición presolar.

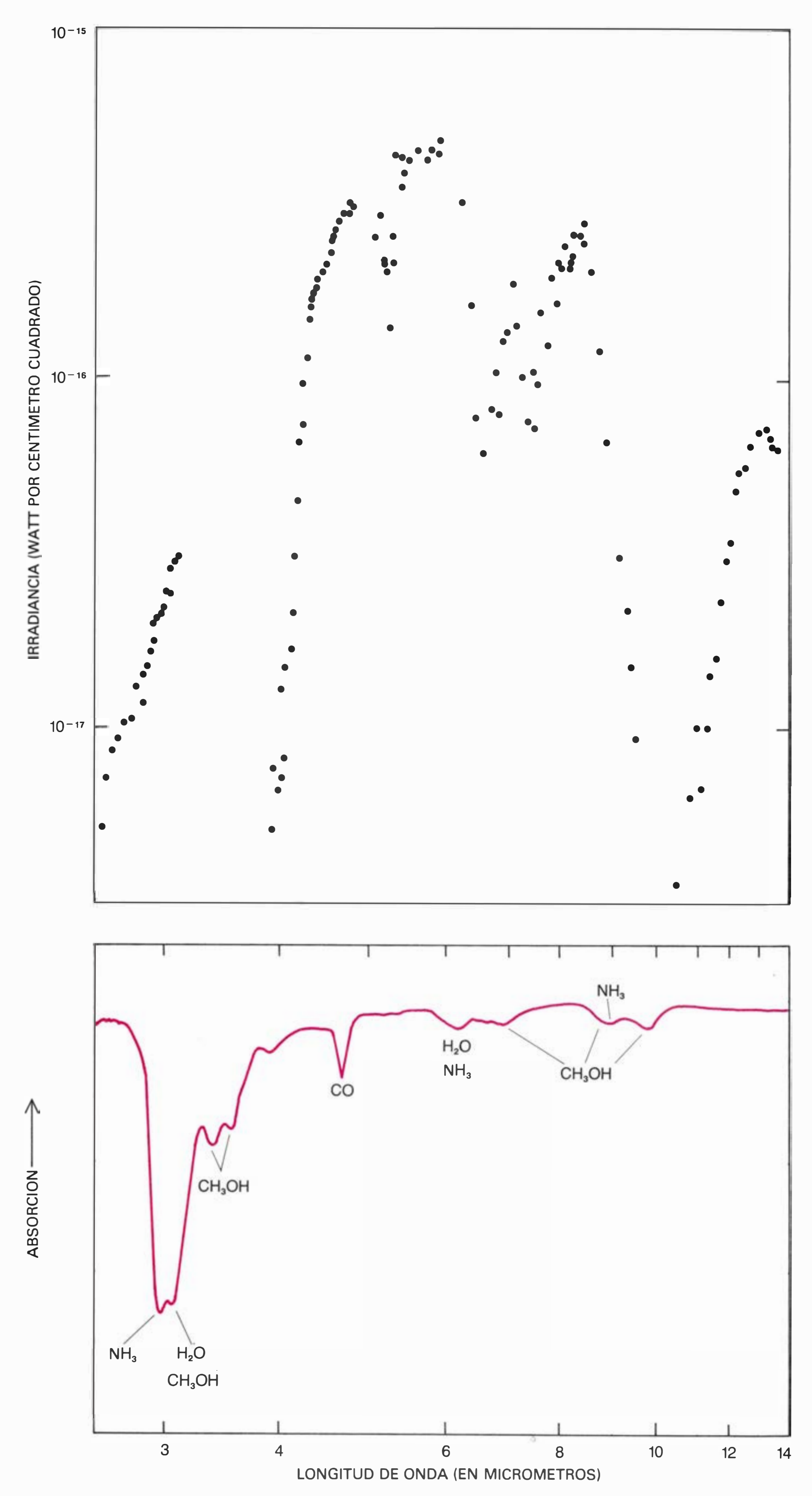
El núcleo de un cometa es un cuerpo sólido de algunos kilómetros de radio, no muchos. Del que, de una manera directa, no sabemos nada. En cambio, se observa el gas y el polvo en la cabeza y la cola del cometa; esto es, la materia que se evapora o se pierde del núcleo de alguna otra manera cuando el cometa pasa cerca del Sol. Tal vez el único punto que apenas se discute es que el núcleo de un cometa incluya una considerable proporción de hielo de agua. En los años cincuenta, Fred L. Whipple, de la Universidad de Harvard, conjeturó que un cometa sería como una bola de nieve sucia. La conjetura que haría yo se basa en la hipotética composición de los granos interestelares que han alcanzado la fase presolar. Creemos que alrededor del 30 por ciento del volumen del núcleo de un cometa es agua; alrededor del 20 por ciento,

materia amarilla; silicatos, en torno al 8 por ciento; sobre el 3 por ciento, carbono puro, y el 40 por ciento restante serían moléculas, entre ellas CO, CO_2 y H_2CO, y radicales, tales como el HCO. En suma, un cometa sería una bola de nieve muy sucia.

El carbono de tal cometa se parecería al carbono de una población bien desarrollada de granos interestelares; así, parte del mismo estaría incorporado en materias volátiles, como CO, CO_2 y H_2CO; otra parte lo encontraríamos en la materia amarilla y parte, por fin, sería carbono puro. Semejante distribución a tres vías da cuenta de lo que Armand H. Delsemme, de la Universidad de Toledo, describe como el "carbono perdido" de los cometas. Con ello quiere significar que la relación de carbono a oxígeno en un cometa (deducida de los espectros de los cometas) es más baja que la misma relación en el sistema solar. ¿Dónde podría hallarse escondido el carbono? Según unos, el carbono detectado en el espectro de un cometa proviene de la provisión de materias volátiles y casi la mitad del carbono del cometa está encerrado en la materia amarilla.

Recientemente, Michael F. A'Hearn, de la Universidad de Maryland en College Park, y Paul D. Feldman, de la Universidad Johns Hopkins, han aportado nuevas pruebas en favor de los gránulos interestelares como constituyentes de los cometas. Han descubierto que en éstos hay moléculas de azufre diatómico (S_2) y afirman que tales moléculas pueden haberse originado solamente en materia sólida bañada por la radiación ultravioleta.

Mientras tanto, los cometas y la Tierra han quedado enlazados de una manera espectacular: se ha sugerido que un ciclo de impactos de cometas en la Tierra de 26 millones de años causó la extinción de especies, entre ellas los dinosaurios. Tales impactos debieron ser frecuentes cuando la Tierra era muy joven, y los cometas depositaron, pues, en ella gránulos interestelares en gran cantidad. Los granos incluirían materia amarilla, es decir, materia orgánica compleja; de esta manea, los granos interestelares pueden haber proporcionado el entorno químico necesario para la evolución prebiótica. Los estudios de laboratorio de los granos interestelares no solamente exploran la física y química del espacio interestelar, sino que también pueden suministrar un enlace entre las observaciones astronómicas realizadas hoy día y la reconstrucción de sucesos en el primitivo sistema solar.



11. ESPECTRO DE UNA FUENTE INFRARROJA, designada por W33 A (*gráfico superior*). Manifiesta varias peculiaridades, de entre las que destacaremos la absorción de radiación infrarroja a longitudes de onda de 3,1, 4,6, 6, 6,8 y 10 micrometros. La fuente es, tal vez, una estrella muy joven, inmersa en una nube molecular. El espectro de laboratorio de H_2O, CO, CH_3OH y NH_3 en una mezcla de seis a tres, a tres y a dos (*gráfico inferior*) da cuenta de cuatro de los máximos. Al parecer, los cuatro en las longitudes de onda inferiores se deben al helado manto de los granos interestelares. La absorción a 10 micrometros puede atribuirse, entre tanto, al núcleo de silicato de los granos interestelares. El espectro de W33 A fue medido por T. B. Soifer y sus colaboradores, adscritos al Instituto de Tecnología de California.

Tuercas y pernos

La idea de que un perno y una tuerca a juego podrían servir de elemento de unión data del siglo XV. Hoy, tuercas, pernos y sus procesos de fabricación siguen en desarrollo activo

Frederick E. Graves

Pensando en cuántas máquinas y otras estructuras tienen elementos de unión, podría creerse que la combinación de tuerca y perno remontan su origen en la lejanía del tiempo y que está tan desarrollado este elemento de unión que nada nuevo podría esperarse de él. La verdad es que la tuerca y el perno datan de mediados del siglo XV, cuando se realizaban a mano y una tuerca apropiada para un perno resultaba improbable que se ajustara a otro. Además, aun reconociendo que la tecnología de la unión por tuerca-perno está bien establecida, se ha beneficiado considerablemente de varios avances recientes. Recordemos los sistemas electrónicos para apriete de la tuerca, los recubrimientos especiales para tuercas y pernos, así como nuevas herramientas para su fabricación, entre otros.

La combinación de tuerca y perno es, por otra parte, un miembro más de la familia de elementos de unión: una gran familia, si contemplamos la diversidad de diseños, tamaños y aplicaciones. Los otros miembros de la familia están representados por cuñas, tornillos y remaches. Me centraré sobre la unión por tuerca-perno, puesto que su historia es el tema central en el desarrollo de partes intercambiables.

Podemos definir el perno como un cilindro externamente roscado, con una cabeza en un extremo. Si el perno no es totalmente roscado de la cabeza a la punta, el diámetro de la parte no roscada viene a coincidir con el diámetro de la cresta de la rosca (el diámetro máximo de la rosca). La forma de la cabeza del perno puede ser cuadrada, hexagonal, redonda, avellanada, elíptica u oval.

La tuerca puede definirse como un prisma de metal, usualmente de sección cuadrada o hexagonal, con un agujero pasante en su centro. El agujero está roscado en su interior, de manera que puede emparejarse con la rosca externa del perno. Varían en su diseño: tuercas de autorretención, tuercas almenadas y tuercas de mariposa. La mayoría de las variantes intentan resolver el problema de la pérdida de tuercas en uniones sometidas a vibración. La tuerca de autoretención incorpora una inserción de nylón u otro material plástico para incrementar la fricción, aumentando su resistencia a la rotación, o puede realizarse con tuercas de roscas desviadas que logran idéntico cometido. Roscas ranuradas o almenadas tienen ranuras en donde puede situarse un pasador que afiance la tuerca en su sitio.

No obstante semejantes ingenios, las oscilaciones o vibraciones provocan, a veces, pérdidas de tuercas. Si se produce el afloje o la pérdida de la tuerca, el perno oscilará y la unión fallará.

Una idea de la importancia fundamental de los elementos de unión se puede adquirir mencionando el número de estos elementos (de todos tipos) que intervienen en una serie de objetos de uso común. Un teléfono tiene 73, un lavavajillas 115, un refrigerador 275, una elevadora transportadora de horquilla 940, un vagón furgón 1200, un torno vertical 1650, un automóvil 3500 y un avión reactor 1,5 millones.

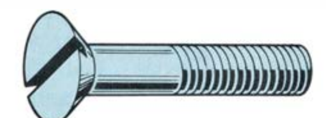
No es fácil mostrar la amplísima gama de tamaños, resistencias y tipos de rosca de pernos que se le ofrecen al ingeniero en sus proyectos. Unos pocos ejemplos de pernos usuales en Estados Unidos puede indicarnos las posibilidades. La norma A-307 de American Society for Testing Materials abarca pernos de acero con bajo contenido en carbono con un rango de diámetros desde 6 hasta cien milímetros para usos generales. Tales pernos deben tener un mínimo de 4200 kilogramos por centímetro cuadrado de tensión admisible. La norma A-325 de dicha organización abarca pernos de acero de alta resistencia con un contenido medio en carbono en los rangos de diámetros desde 13 hasta 38 milímetros; se emplean en uniones de estructuras de acero. La resistencia a tracción del perno será, por lo menos, de 8400 kilogramos por centímetro cuadrado en los diámetros comprendidos entre 13 hasta 25 milímetros y de 7400 kilogramos por centímetro cuadrado en diámetros comprendidos entre 29 hasta 38 milímetros. El grado 8 de la Society of Automotive Engineers incluye pernos desde 6 hasta 38 milímetros en diámetros habitualmente usados en automóviles, maquinaria agrícola y equipos mecánicos donde la resistencia y durabilidad de una unión son decisivas; la resistencia mínima a la tensión es de 10.500 kilogramos por centímetro cuadrado. Normas para pernos utilizados en aviación incluyen tipos con resistencias mínimas a tracción de 13.700 a 18.300 kilogramos por centímetro cuadrado.

Desde el punto de vista económico, la importancia de las uniones se destaca en un reciente informe según el cual, en Gran Bretaña, la unión de componentes contabiliza entre el 20 al 40 por ciento de los costes de fabricación. Semejante proporción podría ser idéntica en los Estados Unidos. La mayor parte de los costos corresponde a la mano de obra. Si bien los costes de los elementos empleados en uniones en un conjunto mecánico no llegan, por término medio, al 5 por ciento del costo del conjunto acabado, en realizar las uniones se invierte algo más que la mitad del tiempo total de producción de un equipo mecánico. El resultado es que el costo de las uniones realizadas entre en el rango de 3 a 10 veces el costo de los elementos de unión.

La industria de fabricación de elementos de unión reviste importancia económica por derecho propio. Alrededor de 625 compañías en los Estados Unidos emplean entre 50.000 y 60.000 trabajadores, que producen en torno a los 250.000 millones de elementos para

PERNOS

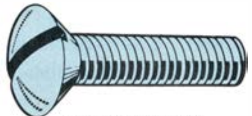

TORNILLO DE MAQUINA
CABEZA REDONDA


TORNILLO DE MAQUINA
CABEZA PLANA


TORNILLO DE MAQUINA
CABEZA HEXAGONAL

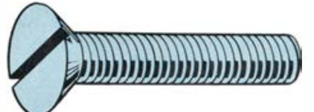

TORNILLO DE MAQUINA
CABEZA HUECA


PERNO SUSTENTADOR


PERNO PUNTAL
CABEZA OVAL

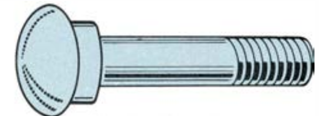

TORNILLO DE MAQUINA
CABEZA FILLISTER


TIRAFONDO


PERNO
CABEZA PLANA


PERNO
CABEZA REDONDA


PERNO DE CARRUAJE


PERNO DE CUELLO
OVAL, CABEZA REDONDA

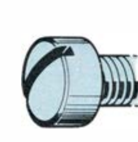
TIPOS DE CABEZAS


REDONDA

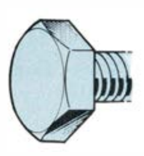

PLANA

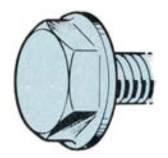

OVAL


PLATO

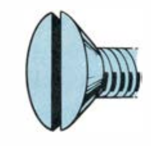

FILLISTER


TRUSS

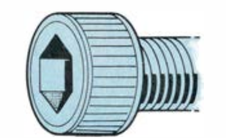

HEXAGONAL

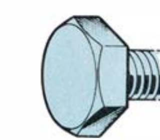

ARANDELADA

ACCIONAMIENTOS


RANURADA


PHILLIPS


HEXAGONAL EXCAVADA


HEXAGONAL


SEGURIDAD

ARANDELAS


INMOVILIZADORA
DE RESORTE


PLANA


DISCO
INMOVILIZADOR


INMOVILIZADORA DE
DENTADO INTERIOR


INMOVILIZADORA DE
DENTADO EXTERIOR

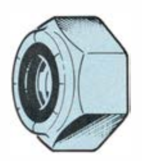

DE PROTECCION

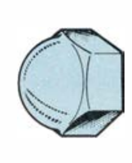

DE ACABADO

TUERCAS

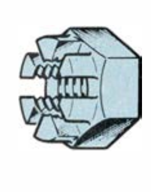

HEXAGONAL


CUADRADA


DE AUTORRETENCION


DE SOMBRERETE


MARIPOSA

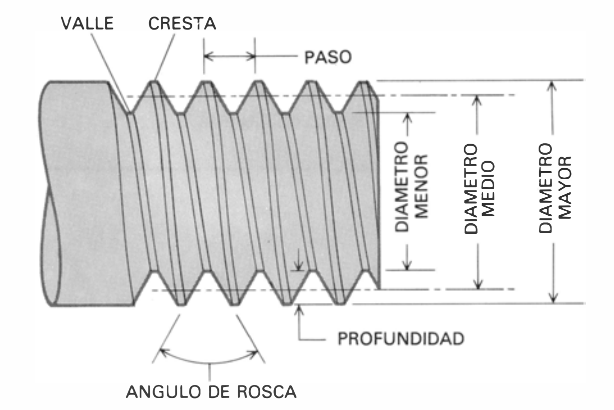

ALMENADA

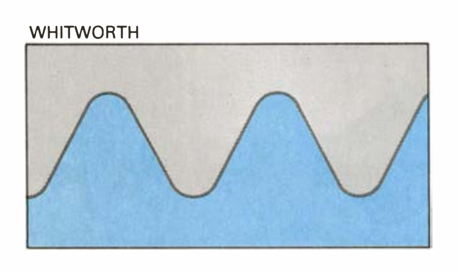
1. GLOSARIO de pernos clasificados por su tipo, forma de cabeza y accionamiento: es decir, el sistema por medio del cual se aplica un par torsor al perno. Se ofrecen varios tipos de arandelas, incluidas las inmovilizadoras que fijan la tuerca una vez apretada. Abajo: distintos tipos de tuercas.

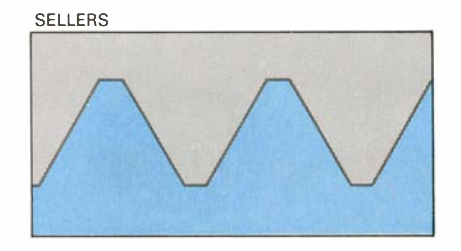
uniones por año. Los consumidores pagan 10.000 millones de dólares por estos elementos de unión y mucho más por su instalación. El costo final de las uniones instaladas se cifra en unos 50.000 millones de dólares.

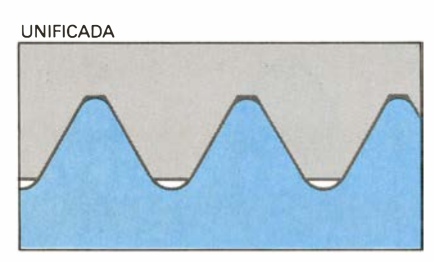
Si bien los pernos roscados datan del siglo XV, los no roscados (de mucha menos aplicación) son más antiguos. La arqueometalurgia muestra que fueron usados por los romanos en las puertas, como quicios y como pasadores, es decir, barras o banquetas con muescas en las que insertaban chavetas para impedir que el pasador se moviera. Los romanos desarrollaron también, según parece, el primer tornillo de madera, trabajado a su vez en bronce y a veces en plata. Sus roscas eran limadas a mano o consistían en un alambre enrollado a la varilla y unido a ella. Todo indica que esta invención desaparece con el Imperio Romano, puesto que el primer registro impreso de un tornillo aparece en un libro fechado en los albores del siglo XV.

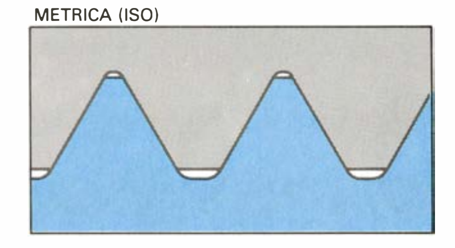
Avanzada ya esa misma centuria, Johann Gutenberg incluyó tornillos entre los elementos de unión utilizados en su imprenta. Anteriormente, los relojeros y armeros usaron también tornillos. (Hay un ejemplar de armadura ecuestre española fechada en 1614 que contiene numerosos agujeros roscados para afianzar los guiones.) Cuadernos de Leonardo da Vinci, de finales del siglo XV y principios del XVI, recogen varios diseños para máquinas de fabricar roscas, pero la primera máquina utilizable en la práctica con este propósito fue inventada en 1568 por Jacques Besson, un matemático francés. Antes de terminar el siglo XVII, los tornillos eran ya componentes comunes en las armas de fuego.

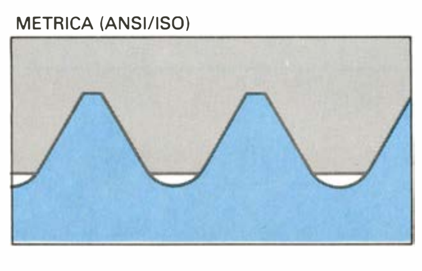












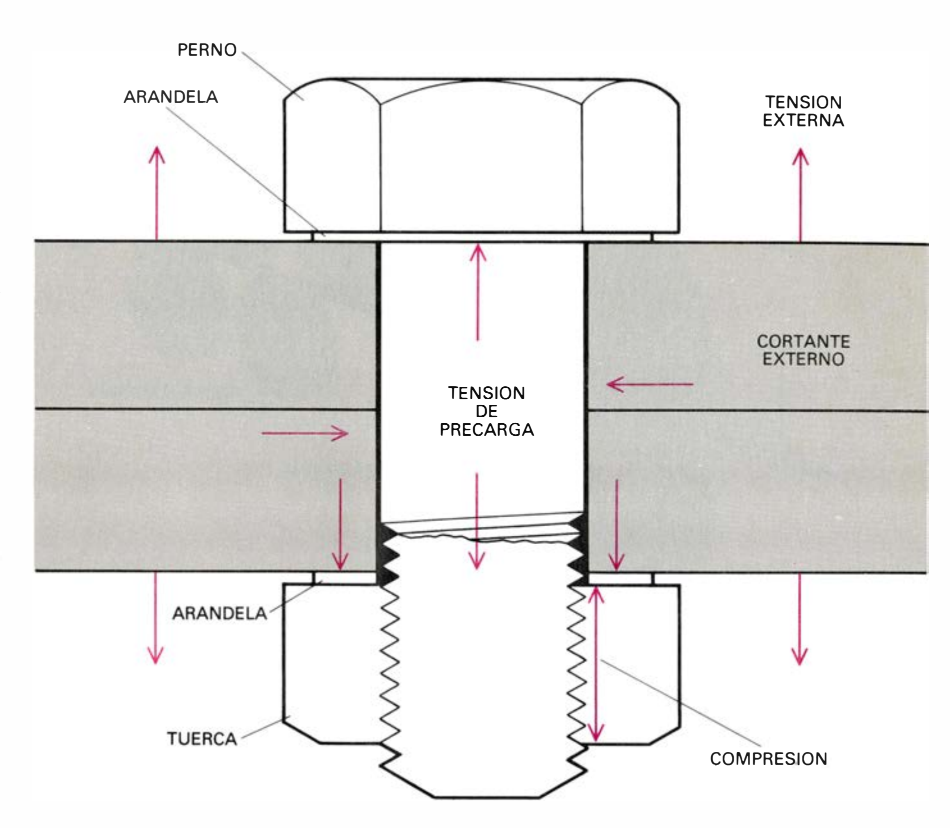
2. ROSCAS DE PERNO, con su nomenclatura y varias formas normalizadas. La rosca Whitworth, diseñada por Joseph Whitworth en 1841, se convirtió en norma británica; la rosca Sellers, diseñada por William Sellers en 1864, se constituyó en norma norteamericana. El Sistema Unificado, que incorporó características de las normas Whitworth y Sellers, fue adoptado por los Estados Unidos, Canadá e Inglaterra en 1948. La norma métrica ideada por la Internacional Organization for Standardization (ISO) se adoptó en 1964 por las naciones que empleaban el sistema métrico. En 1971 el Industrial Fasteners Institute propuso el Sistema Métrico Optimo, que mejoraba la norma ISO. Del mismo se originó el sistema de la American National Standards Institute/ISO, considerado la norma métrica internacional.

Con estos desarrollos de los pernos y de las roscas realizadas a mano, la tuerca estaba pronta a aparecer, así como la idea de añadir una tuerca a este perno. La primera referencia impresa de una tuerca roscada apareció a finales del siglo XVI y comienzos del XVII. Como los primeros tornillos, las primeras tuercas eran realizadas a mano en forma muy laboriosa.

Por los datos de que se dispone, fue en los primeros años del siglo XVII cuando se añadieron tuercas a los tornillos fabricados en la época, de contorno recto y de extremo embotado; mostraban, pues, más parecido a los modernos pernos que a los puntiagudos tornillos actuales. Un libro inglés de 1611 habla de una "tuerca para un tornillo". El tornillo ya podría calificarse como un perno. Pero que la rosca de una tuerca se acomodase a la rosca de su perno particular era cuestión de suerte; cuando semejante acuerdo se encontraba, la tuerca y el perno se separaban juntos hasta su instalación en una máquina o una estructura.

Como puede suponerse, fue con el advenimiento de la revolución industrial cuando los pernos y las tuercas adquieren un uso común como elementos de unión. Esta época de grandes realizaciones comienza con la invención por James Watt de la máquina de vapor en 1765. Los fabricantes de estas primeras máquinas y de los productos que rendían comprendieron en seguida que las uniones roscadas resultaban decisivas en los diseños eficientes de conjuntos mecánicos, por la facilidad en los montajes de estas máquinas y productos, así como por su seguridad. Muchas invenciones famosas son meras extensiones de la idea de las uniones roscadas. Citemos, por ejemplo, la máquina de hilar de James Hargreaves y la desmotadora del algodón de Eli Whitney.

Whitney preparó, en 1801, el camino para el siguiente concepto fundamental: la intercambiabilidad de los componentes. En ese año se presentó ante un grupo de oficiales en Washington, entre los que se contaban el presidente Adams y el vicepresidente Jefferson. Formó varias pilas con componentes de mosquetes sobre una mesa; cada pila contenía 10 componentes idénticos. Tomando un componente al azar de cada pila, no tardó en montar un mosquete que funcionaba. La idea de la intercambiabilidad afianzó de este modo su solidez e intervino en otros inventos; así: el revólver de Samuel Colt, el martillo pilón de James Nasmyth, la segadora de Cyrus McCormick, el arado de acero de John Deere, la producción en serie de relojes de Eli Terry y la máquina de coser de Elias Howe.



3. FUERZAS que se desarrollan en una unión. Forman esta unión una tuerca, un perno y una arandela. En teoría, el apriete debería realizarse hasta el punto de fluencia, cuando el metal comienza a deformarse.

Un problema que persistió hasta los inicios del siglo XIX fue la falta de regularidad en las roscas de tuercas y pernos. Hasta finales del siglo XVIII, la técnica habitual para conformar roscas en gran número, labor que realizaban los herreros, era golpear un perno sin rosca con un punzón especial, que actuaba como herramienta de forma. Las roscas más pequeñas se cortaban en un torno primitivo. El operario solía sostener la herramienta de corte: resultaba, pues, prácticamente imposible conseguir roscas regulares.

Hacia 1800, el torno se mejoró con la introducción del carro y el accionamiento; la rosca se desplazaba así uniformemente por medio de un tornillo de avance –el tornillo que movía la herramienta de corte– y podía reproducirse con seguridad; no se disponía, sin embargo, de un sistema que relacionara el número de filetes de rosca con el diámetro del tornillo. Como Nasmyth anotó: "pernos y tuercas correspondientes deberían fabricarse compatibles entre sí. Entremezclarlos... supondría caer en dudas y costes sinfín, en faltas de rendimiento y en confusiones, en particular cuando se tratara de reparar partes de máquinas complejas".

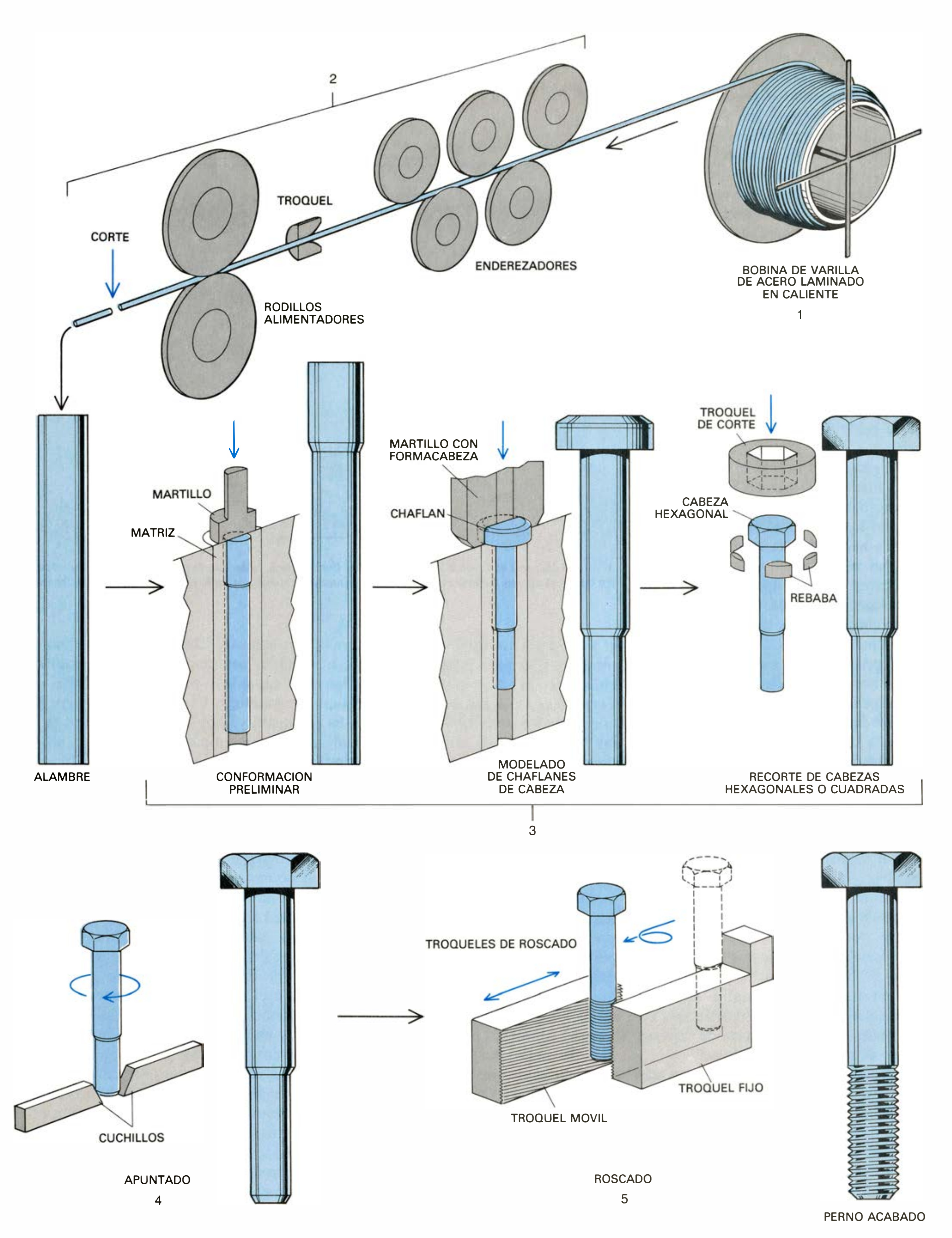
El hombre que dio un giro a la situación fue el inventor inglés Henry Maudslay. En 1800 fabricó el primer torno que permitió al operario hacer tornillos de cualquier paso y diámetro de los que tuviera necesidad. (Paso es la distancia entre la cresta de un hilo de rosca a la cresta del siguiente hilo. El diámetro mayor se mide de la cresta de un filete de rosca a la cresta del filete opuesto del perno; el diámetro menor, desde el valle de un filete de rosca hasta el valle del filete opuesto.) Su contemporáneo Charles Holtzapffel escribió, en su trabajo en cinco volúmenes sobre *Turning and Mechanical Manipulation*, que entre los años 1800 y 1810 Maudslay "cambió casi de arriba abajo los antiguos, imperfectos y azarosos métodos de fabricación de tornillos... en el modo moderno, exacto y científico, ahora seguido por la mayoría de los ingenieros".

El torno de fabricación de tornillos constituyó, durante muchos años, el medio principal de producción de roscas. La práctica normal hoy consiste en presionar matrices roscadas contra un perno sin rosca, haciéndolo girar. La diferencia crucial es que el procedimiento de corte con torno arranca material, mientras que el estampado por troquel conforma la superficie sin pérdida de material.

A mediados del siglo XIX, William E. Ward, de Port Chester, Nueva York, desarrolló mecanismos para la forja de tuercas y pernos. En su procedimiento, una porción de barra se calentaba hasta 870 grados Celsius y se mantenía dentro de troqueles de forma. Posteriormente, Ward desarrolló otros mecanismos para procesado en frío. En este proceso, no se sometía la barra a calentamiento. El troquel debía ser fuerte y la maquinaria empleada en este proceso, potente. El conformado en frío originaba tolerancias más estrechas que en caliente y producía elementos más resistentes. El conformado en frío constituye hoy el método básico de fabricación en masa de tuercas, pernos y tornillos.

En las postrimerías del siglo XIX, la fabricación de uniones en serie progresó, de una manera gradual, desde la manipulación de la barra hacia la conformación en frío continua a partir de bobinas o alambres. La bobina alimenta una máquina de fabricación de pernos. La máquina corta varillas de longitud adecuada. La varilla atraviesa luego una serie de troqueles y se convierte en alambre, sobre el que se trabaja el enroscado. El proceso de conformación continua en frío es seguido hoy en la mayoría de las industrias de pernos y tuercas.

La capacidad de fabricación de roscas regulares no garantizaba todavía su uniformidad, puesto que cada fabricante realizaba su propio tipo de rosca. Urgía, pues, una normativa nacional e internacional. En el Reino Unido el primer movimiento significativo en esta dirección llegó en 1841, cuando Joseph Whitworth presentó al Instituto de Ingenieros Civiles *A Uniform System of Screw-Threads* ("Un sistema uniforme de roscas de tornillos").



4. FABRICACION DE UN PERNO por el proceso de conformado en frío. Se comienza a partir de una bobina de varilla de acero laminado en caliente (*1*). Los rodillos enderezan la varilla; en cuanto un troquel especial la reduzca al diámetro deseado, se corta (*2*). El alambre resultante se introduce en una máquina que realiza tres operaciones: conformación inicial, modelado de una cabeza circular achaflanada inicial y finalmente recorte de esta cabeza a su forma cuadrada o hexagonal (*3*). Se apunta el perno (*4*) y se rosca entre dos troqueles, uno fijo y otro móvil. Las tuercas se fabrican de modo similar.

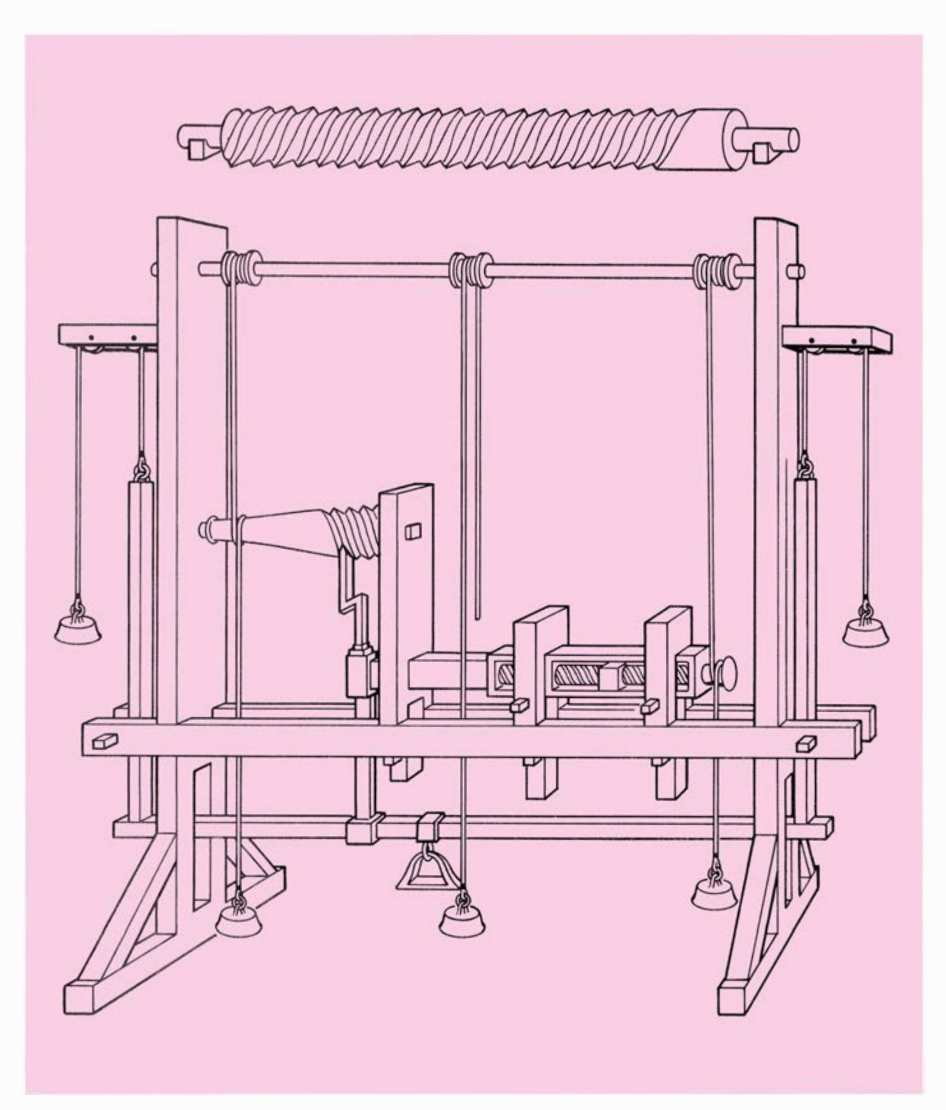
Whitworth propuso que, para tornillos y pernos de un tamaño dado, las roscas debían ser idénticas en paso, profundidad y forma. Recomendó un ángulo de 55 grados como ángulo de rosca. El número de hilos de rosca por centímetro debería especificarse para cada diámetro de tornillo o perno. El filete de rosca estaría redondeado en la cresta y en el valle en un sexto de la profundidad de rosca. Hacia 1881, el sistema Whitworth se generalizó como la norma británica.

En Estados Unidos, el movimiento hacia la normalización comenzó en 1864. William Sellers, fabricante de herramientas de Filadelfia, convenció al Franklin Institute de aquella ciudad que formara un comité al objeto de buscar y establecer una norma nacional. Sellers tenía varias reservas para con el sistema Whitworth. Asegurando que el ángulo de 55 grados de corte era difícil de medir, propuso el de 60 grados, que, añadía, originaba además filetes de rosca más resistentes. Sostenía también que las roscas redondeadas de las crestas del sistema Whitworth producían un encaje incierto y débil entre perno y tuerca. Propuso roscas con crestas y valles planos.

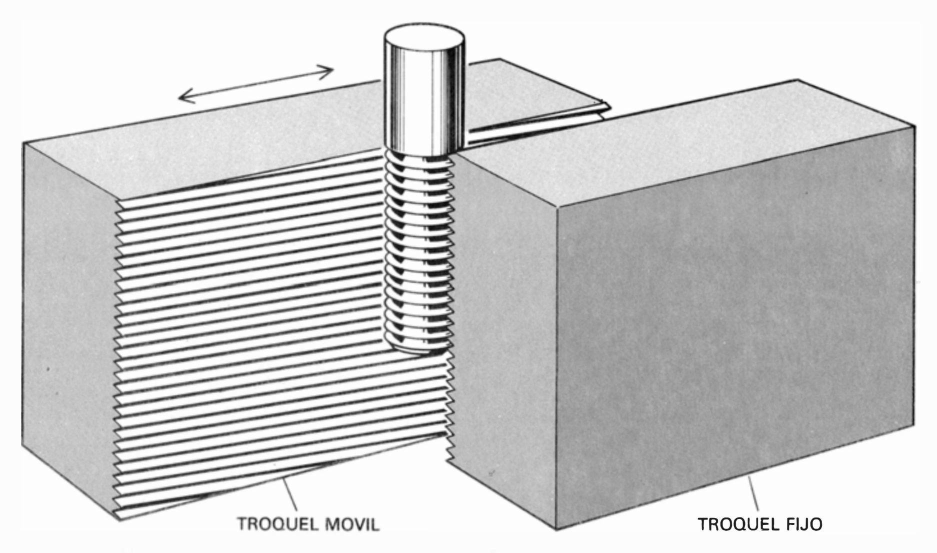
En su informe final, el Franklin Institute adoptó el sistema de Sellers. Recomendaba como norma nacional que "las roscas de tornillos tuvieran cantos rectos y un ángulo entre cantos de 60 grados, terminado en superficies planas en la cresta y valles iguales a un octavo del paso". A finales del siglo, el sistema Sellers se convirtió en la norma de los Estados Unidos y gran parte de Europa.

La incompatibilidad entre los sistemas Whitworth y Sellers planteó serias dificultades durante la primera y la segunda guerra mundial, cuando las fuerzas armadas de los Estados Unidos y Gran Bretaña necesitaron intercambiar material de campaña. Desde 1918 y continuando intermitentemente hasta 1948, ambos países buscaron caminos que reconciliaran los respectivos sistemas. En una conferencia celebrada en Washington en 1948, los Estados Unidos, Canadá, y Reino Unido adoptaron el Sistema Unificado de Rosca que reunía características de los sistemas Whitworth y Sellers. En la adopción de pulgadas como unidad en roscas intervino de una manera decisiva el Instituto de Uniones Industriales, el mayor fabricante de uniones de los Estados Unidos.

En el mismo año, la Organización In-



5. TORNO PARA ROSCADOS de 1568, de Jacques Besson. Fue la primera máquina útil para este propósito y contribuyó significativamente a la introducción de uniones por tuerca-perno. El operador gobierna la máquina por medio del estribo de la parte inferior. La pieza que se trabaja: el tornillo cónico de la izquierda; se está cortando su rosca por una herramienta movida por el tornillo guía de la derecha. El movimiento de la herramienta y la rotación de la pieza que se trabaja se coordinan por las poleas.



6. PROCESO MODERNO DE ROSCADO. No se trata, en absoluto, de una operación de corte que pierda material, sino de un remodelado de la superficie por troqueles. En la máquina que se emplee, se presiona el alambre de perno entre dos troqueles, uno fijo y el otro móvil, que origina el giro del perno.

ternacional para Normalización (ISO) empezó a trabajar sobre un sistema de rosca para tornillo normalizado de aplicación mundial. Cuando los trabajos terminaron en 1964, el sistema adoptado en la conferencia internacional celebrada en Nueva Delhi consistía en realidad en dos sistemas: el sistema ISO en pulgadas para roscas de tornillos (igual al Sistema Unificado) y el sistema ISO métrico para roscas de tornillos, que se creó para sustituir los múltiples y distintos sistemas métricos nacionales que existían.

Puesto que los elementos de unión realizados con el sistema métrico resultaban inferiores a los realizados con el sistema ISO en pulgadas, el Instituto de Uniones Industriales recomendó, en 1970, que se diseñara un sistema métrico mejor. En 1971 el grupo propuso el Sistema Métrico Optimo para Uniones. Además de otras cosas, se pedía un perfil de rosca, basado en un tipo que llegó a normalizarse en el campo aerospacial y como uniones de mejor resistencia a la fatiga del metal. Esta propuesta condujo a un sistema similar, que es ahora la norma métrica internacional: el sistema ANSI/ISO. (ANSI son las iniciales de American National Standards Institute.)

Otras organizaciones se preocupan por hallar normas de uniones, a menudo relacionadas con la que les resulta más apropiada para una industria particular. De este tenor son la American Society for Testing and Materials, American National Standards Institute (ANSI), Society of Automotive Engineers (SAE) y otras. En conjunto su radio de actividad abarca unas 8000 normas de uniones, donde entran características tales como clase de materiales, configuraciones, dimensiones, tolerancias y otras propiedades mecánicas. Si agregamos uniones especiales y propias, acabados superficiales y recubrimientos y todas las combinaciones de diámetros y longitudes, el número de uniones diferentes sobrepasan el par de millones.

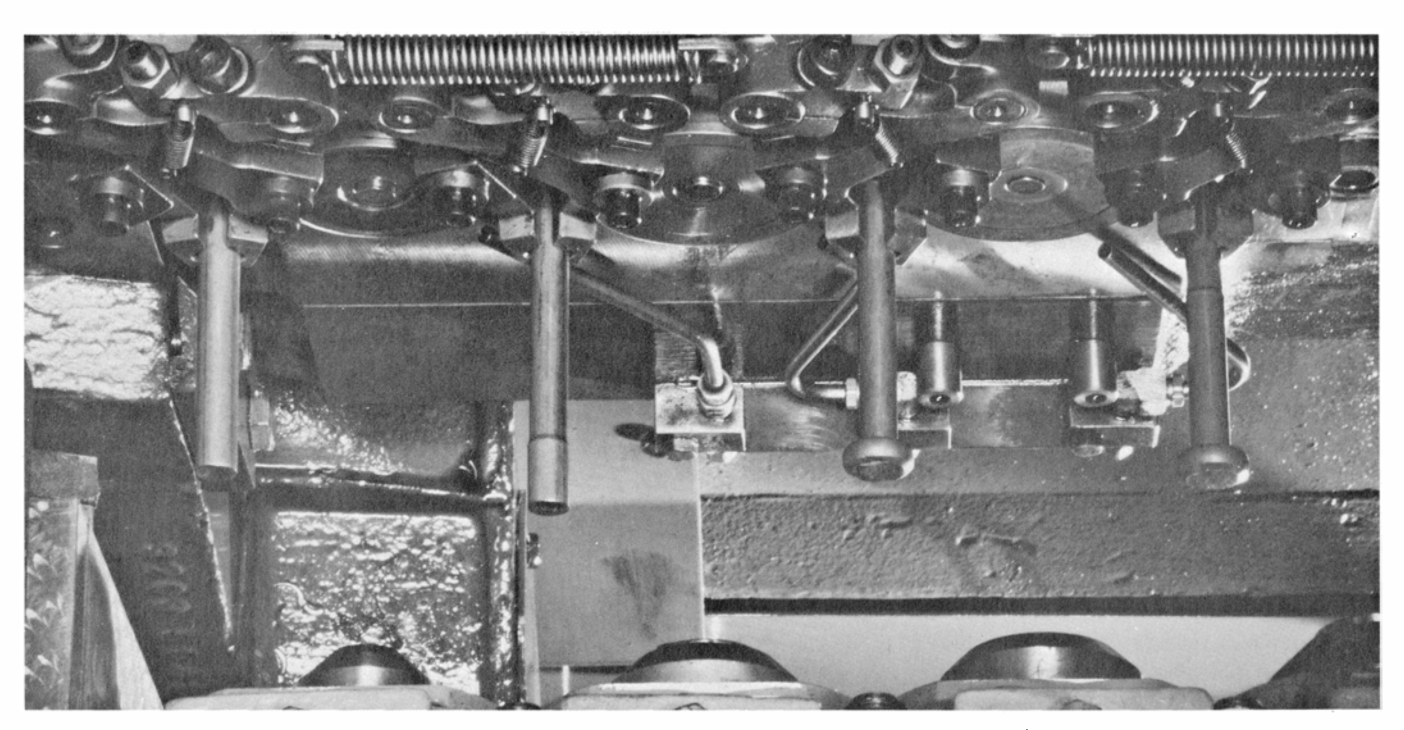
Con tantas para elegir, el fabricante haría bien en estar prevenido contra la proliferación de tipos y estilos de uniones. A menos que tenga cuidado en el momento del diseño de una parte de un equipo, la proliferación de uniones puede ocasionar unos significativos aumentos en los costes.

Supóngase que se comienza a diseñar una máquina que requerirá la instalación de un tornillo autorroscante en cierta etapa del montaje. Supóngase también que sólo se necesitará una combinación de diámetro y longitud. Para esta aplicación el diseñador dispuso de las siguientes variantes en unas normas para tornillos autorroscantes: nueve formas de rosca, seis tipos de cabeza, tres tipos de ranura de cabeza y cuatro tipos de acabado superficial. El número total de variantes era 648. Es fácil imaginar así que una compañía con varios diseñadores y delineantes puede acabar con un aluvión de uniones, cada una con un número de código y cajas repletas en almacén.

Un ejemplo sencillo nos pondrá de manifiesto que el problema no se atenúa. Se pide una combinación de perno, tuerca y arandela. Comprobando las posibilidades de las uniones normalizadas, el ingeniero encontró que para un perno de una longitud y diámetro dados, podía escoger entre tres tipos de cabeza, cuatro especificaciones de resistencia, dos pasos de rosca, tres acabados superficiales y al menos dos tipos de arandelas, así como dos tipos de tuercas. El número de combinaciones posibles suma hasta 288.

Se ha establecido en más de 2000 dólares anuales el costo de mantener un único tipo en inventario. El tipo de unión debe almacenarse en la base de datos de un ordenador de donde puede recuperarse la información. Hay que disponer de un inventario del número de variedades de un determinado tipo de unión, reservar espacio para su almacenaje y contar con un sistema para encontrarlo en el almacén. Todo esto para cada perno, tuerca, arandela u otra clase de uniones que se emplean en fabricación de máquinas.

Muchos fabricantes tienden a conceder una importancia a las uniones menor que la que dan a otros componentes de sus operaciones de ensamblaje. Está muy extendido el prejuicio de que las tuercas y pernos son elementos normalizados, baratos y siempre lis-



7. MAQUINA DE CONFORMADO DE PERNOS, fotografiada en el momento en que el alambre de perno va a ser trabajado por troqueles que ascienden. Se distinguen cuatro pernos en proceso de fabricación. La máquina, de la National Machinery Company, produce más de 100 pernos por minuto.

tos para sacarlos del cajón en el momento en que se precisan. Y así vemos que, salvo contadas excepciones (los fabricantes de automóviles y maquinaria agrícola entre ellos), se derrocha el dinero con un aluvión de uniones innecesarias y el uso de uniones que exceden los requerimientos del caso en resistencia, material, tipo de rosca y recubrimientos. Mi experiencia muestra que algunos fabricantes gastan un millón de dólares por año en uniones, del que podrían economizar el 15 por ciento, de prestar una mayor atención a esta parte de la operación.

La idea de que prestar atención a las uniones valía la pena ha ido abriéndose camino en los últimos años hasta configurar una nueva disciplina ingenieril, que estudia las aplicaciones de las uniones. Cuando proyecta una máquina, el ingeniero considera el diseño de las uniones que necesitará, el tipo de uniones que deberá escoger del inmenso número disponible y los métodos de su instalación. El resultado es un gran incremento en la confianza y seguridad de las máquinas.

Junto con una actitud hasta cierto punto despreocupada hacia las tuercas y los pernos, en muchas compañías es notoria la falta de conocimiento sobre los métodos idóneos de apriete de una unión tuerca-perno. Para una unión eficiente, el apriete debe acercarse a la no muy afortunadamente llamada región de fluencia. El término se refiere a la etapa en que comienza la deformación permanente del metal (específicamente, una elongación del perno). Queda lejos del punto de rotura y resulta ser la máxima presión de apriete.

Debido, en general, a que este proceso suele realizarse a mano, las uniones mayores (16 milímetros o más de diámetro) tienden a ser subapretadas, en tanto que las más pequeñas tienden a sobreapretarse. Existen múltiples llaves especiales así como técnicas de apriete para solucionar este problema; cuando se trata de uniones críticas por razón de sus altas prestaciones o seguridad, la novedad del control electrónico constituye un hallazgo que encuentra cada vez más aplicaciones. En este sistema, un microprocesador regula la operación de apriete (que se deja hoy ya en manos de un robot) empleando un par torsor y señales de ángulo para abrir una llave eléctrica en el punto de fluencia.

El estudio atento del apriete permite colocar una unión más pequeña o de menor calidad en un determinado punto y ahorrar hasta 10 centavos de dólar por unión. Un fabricante que monte un millón de máquinas anualmente puede reducir sus costos en 100.000 dólares por año. La unión optimizada significa también que no se habrá de enviar tantas máquinas al fabricante para su reparación; y ello supone otro capítulo en que ahorrar.

Muchas uniones han de operar sometidas a condiciones ambientales corrosivas, a altas temperaturas o a ambas. Lo normal ha venido siendo que a una unión que trabaje en esas condiciones se le aplique una capa consistente en fosfato y aceite o un óxido negro. Si se requiere una elevada resistencia a la corrosión se aplica cinc o cadmio por electrochapados. Todas estas técnicas originan pérdidas de cualidades mecánicas. El cinc y el cadmio, por ejemplo, pueden causar fragilidad en el metal.

Se ha avanzado en el sentido de adaptar a las uniones una técnica de recubrimiento con aluminio que se utiliza en la industria aeroespacial para componentes que trabajan en condiciones extremas, como en las turbinas. El aluminio se dispersa en un medio de fosfato o cromato. La capa de aluminio incrementa la facilidad de la unión para resistir altas temperaturas, así como muchos líquidos inorgánicos y orgánicos.

Otro adelanto lo tenemos en el recubrimiento para herramientas empleadas en el conformado en frío de uniones. Estas herramientas se fabrican con materiales duros y resistentes, pero se desgastan muy pronto. Hasta el punto de que lo más caro en la fabricación de uniones es el costo de remplazar la herramienta.

Los esfuerzos desplegados para prolongar la vida de las herramientas de conformado cubren distintos tratamientos superficiales: cementación, nituración y electrochapado. Las herramientas se mejoraron con inserciones de carburo de titanio o nitruro de titanio en los lugares de severo desgaste. La nueva técnica consiste en aplicar una delgada capa de carburo o nitruro de titanio por deposición química o física a partir de vapor. Las capas no son sólo extremadamente duras; tienen, además, propiedades lubricantes que disminuyen la fricción entre la herramienta y el elemento. El tratamiento permite que la mayoría de las herramientas sometidas a estos procesos de recubrimiento perduren tres o cinco veces más que las antiguas.

Juegos de ordenador

Del computador a spaghetti y otros artilugios analógicos

A. K. Dewdney

En nuestros días, tan pronto se mencionan las computadoras, todos invariablemente pensamos en las máquinas digitales. Perdidas en el olvido están sus parientes analógicas, ingenios electromecánicos que en otro tiempo dominaron el mundo de la computación. Antes de la Segunda Guerra Mundial, los principales laboratorios se valían de refinadas máquinas analógicas, de nombres tales como intégrafo continuo o analizador diferencial. Aquellas máquinas sacaban partido de analogías mecánicas o eléctricas de las fórmulas matemáticas; las variables e incógnitas podían venir representadas por una cierta magnitud física: una tensión eléctrica o el ángulo girado por un eje, por ejemplo. Las computadoras analógicas se aplicaban al análisis de las ecuaciones originadas en muchos campos: balística, aerodinámica, redes de alta tensión, etcétera.

Las máquinas analógicas no pudieron competir con los ordenadores digitales construidos durante la Segunda Guerra Mundial e inmediatamente después. Las computadoras analógicas solamente parecían capaces de resolver ciertas categorías de problemas, principalmente los definidos mediante ecuaciones diferenciales e, incluso en tales casos, con precisión que no alcanzaba a ser perfecta. Por el contrario, pronto se alcanzó a ver que los ordenadores digitales podían programarse para resolver una variedad infinita de problemas, con muy alta precisión. Paulatinamente, las máquinas analógicas fueron quedando relegadas; a mediados del decenio de 1960 apenas se las mencionaba en los libros sobre computación.

Aunque las descendientes de las grandes máquinas analógicas de hace 50 o 60 años siguen utilizándose todavía en unos pocos laboratorios (donde no dejan de desarrollarlas y perfeccionarlas), nadie mantiene hoy esperanzas fundadas de que las máquinas analógicas vuelvan algún día a tomar la delantera. No puede darse marcha atrás en la revolución digital. Es también cierto, sin embargo, que a veces la revolución puede deslumbrarnos, impidiéndonos ver los atractivos de situaciones anteriores.

1. *El computador a spaghetti ordena números en sucesión decreciente*

El rasgo más delicioso de un cálculo analógico está en la impresión de recibir algo a cambio de nada. Problemas que exigirían horas de cálculo manual (e incluso con ordenador digital) quedan resueltos rápidamente sin más que observar cómo un sistema físico alcanza rápidamente su equilibrio. En los casos más llamativos el proceso parece desarrollarse instantáneamente, y tenemos un artilugio que computa.

Fijémonos por ejemplo en la computadora DAS (iniciales de Dispositivo Analógico a Spaghetti). En la configuración que tuve ocasión de ensayar, el dispositivo era capaz de ordenar hasta 700 números según valor decreciente. La ordenación de listas de números es tarea frecuente en computación digital, por lo que los algoritmos encargados de ejecutarla han sido perfeccionados y depurados al máximo. No obstante, el tiempo necesario para ordenar listas de números crece todavía algo más rápidamente que el tamaño de la lista. En la computadora DAS es preciso invertir algún tiempo en preparar la máquina para operar sobre la lista numérica concreta de que se trate, y otro más en leer los resultados; el proceso de ordenación, en cambio, parece ser instantáneo.

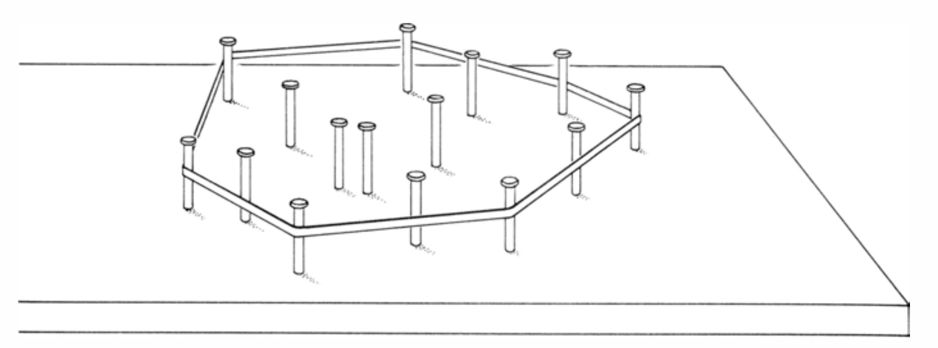
He aquí cómo funciona el dispositivo DAS. Para cada uno de los números de la serie a ordenar se corta una varilla de spaghetti crudo, a la longitud del número. Como es natural, habrá que adoptar unidades de medida apropiadas. Seguidamente, se cogen con una mano todas las varillas así cortadas y, sujetando flojamente la gavilla en posición vertical, damos con uno de sus extremos un golpe seco contra la mesa. El momento cinético que cada una de las varillas adquiere asegura que todas ellas tengan

uno de sus extremos en contacto con la superficie de la mesa. Para obtener la serie ordenada a partir del manojo en reposo hay que ir sacando del haz la varilla más alta, después, la más alta de las restantes, y así sucesivamente hasta agotarlas. Al tiempo de extraer cada varilla se la mide y toma nota del número correspondiente.

Conviene distinguir tres fases en el proceso de ordenación, que llamaré fase preoperatoria, fase analógica y fase postoperatoria. En la fase previa, o preoperatoria, se miden y recortan a la medida los spaghetti; en la fase analógica, un mecanismo sencillo se encarga de la ordenación de las varillas de pasta; en la fase postoperatoria se van extrayendo las varillas de una en una, generando así la serie ordenada. Todos los artilugios descritos en este artículo exigen fases pre- y postoperatorias. Lo mismo ocurría, en efecto, en las antiguas máquinas de cálculo analógico; mucho tiempo y esfuerzo se han invertido en investigación y desarrollo de métodos para acelerar estas fases.

Sin duda habrá quien diga “Tal vez DAS pueda ordenar 700 números, pero ¿qué pasaría si fueran 7000, o 7 millones?” El único tratamiento razonable que merecen cuestiones como ésta es mencionar la existencia de SUPERDAS, ingenio obtenido por modificación de un carretón con elevador de horquilla, capaz de sujetar de una vez 7 millones de spaghetti crudos, recortados a la longitud exacta, y darles un golpe seco por una de sus puntas contra una pared de ladrillo.

Otros habrá que digan: “Yo soy capaz de ordenar 700 números más rápidamente que pueda hacerlo DAS. Me bastan lápiz y papel, ¡y habré terminado antes de que se haya cortado los spaghetti!” Por desgracia para DAS, lo más probable es que esta objeción resulte verdadera. Supongamos que para leer un número de la serie desordenada, medir en la varilla de spaghetti la longitud precisa, cortarla e insertarla en el haz, haga falta un minuto. Supongamos, además, que se tarden 10 segundos en retirar y medir cada varilla tras haberse ejecutado la operación de golpearlos contra la mesa (un segundo). En tales hipótesis, DAS precisaría más de 13 horas para ordenar los 700 números, invirtiendo todo ese tiempo, menos un segundo, en las fases preoperatoria y postoperatoria. El humano, por su parte, ordena la lista inspeccionándola, seleccionando el número máximo, anotándolo y tachándolo de la lista. Para examinar cada número nece-



2. Determinación de la envoltura convexa de un conjunto de puntos del plano

TIEMPO DE COMPUTO

EXPONENCIAL

CUADRATICO

LOGARITMICO

LINEAL

CONSTANTE

TAMAÑO DEL PROBLEMA

3. Posibles funciones que relacionan tiempo de cómputo y tamaño del problema

sita una décima de segundo, y por tanto, para ordenar la serie completa tardará (700 + 699 + ... + 2 + 1)/10 segundos, que resultan ser 175 minutos y 15 segundos. ¡Victoria clara del aspirante! ¿Quedará destruida la maravillosa e instantánea cualidad de la fase analógica por las fastidiosas etapas pre- y postoperatorias? No del todo.

Supongamos que el humano, engreído, reta al SUPERDAS a ordenar 7 millones de números. Mientras que SUPERDAS tardaría unos 15 años, ¡el humano no podría terminar en menos de 74.000! La superioridad de SUPERDAS queda explicada al examinar cómo depende del cómputo del tamaño del problema el tiempo de ejecución. El impacto de los spaghetti contra la pared es una operación que requiere tiempo constante: es intrínsecamente independiente del tamaño del haz. El tiempo requerido por las fases preoperatoria y postoperatoria crece "linealmente": su duración es directamente proporcional al tamaño de la lista de números. La tarea que ha de afrontar la persona se agranda mucho más rápidamente, porque ha de inspeccionar todos los números que aún queden en la lista cada vez que la repasa. Si bien la serie va reduciéndose progresivamente, el número total de inspecciones elementales crece según el cuadrado del tamaño de la lista; se dice que el procedimiento de ordenación es de duración cuadrática. Los procedimientos de duración lineal acabarán, antes o después, por vencer a los de duración cuadrática.

A pesar de sus larguísimas fases de preparación de datos y lectura de los resultados, una máquina DAS suficientemente grande acabaría venciendo a cualquiera de los modernos ordenadores digitales, aunque éste utilizase los más veloces algoritmos de ordenación. Requieren éstos un número de pasos elementales proporcional a $n \log n$, donde n es el tamaño de la lista y log es la función logarítmica. Aunque en teoría los procesos de duraciones linealmente crecientes sean preferibles, hace temblar el mero pensamiento del tamaño que habría de tener el dispositivo spaghetti para vencer en semejante confrontación. Tal máquina, que llamaremos SUPERDUPERSAG, tendría que ensamblarse en el espacio extraterrestre y dar topetazos contra la Luna.

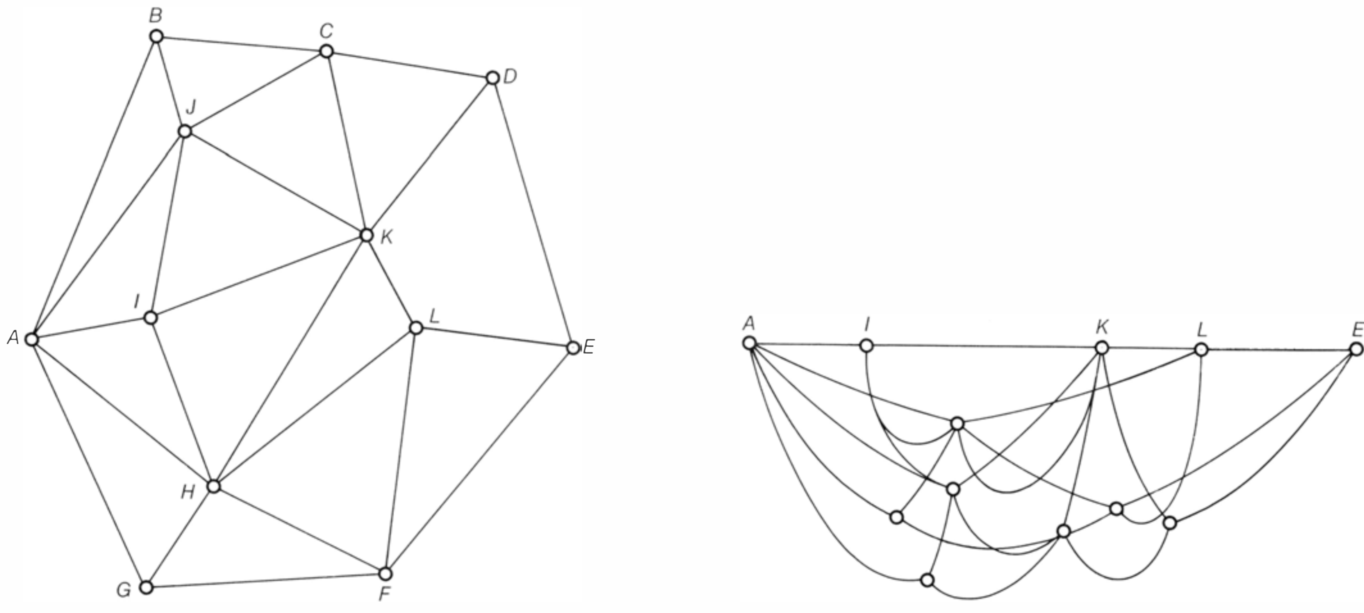
El siguiente artefacto de mi colección sirve para calcular la envoltura o cápsula convexa de un conjunto de n puntos del plano. La envoltura convexa es, sencillamente, la mínima región convexa capaz de contenerlos, y queda totalmente definida por su contorno, un polígono, cada uno de cuyos vértices es alguno de los puntos dados. El dispositivo que calcula el contorno consiste en un tablero grande, algunos clavos y una banda elástica de goma. Lo hemos bautizado CABE (Computadora Analógica de Banda Elástica). Para ponerlo a punto basta clavar en el tablero n puntas de carpintero, en las posiciones correspondientes a los n puntos dados del plano. Se coge después la banda elástica y, estirándola en forma de anillo grande, rodeamos con ella el grupo de clavos. Al soltarla, la anilla de goma se ciñe instantáneamente en torno a los clavos, determinando el contorno poligonal de la cápsula convexa.

Como antes, la operación analógica –soltar la banda de goma– va precedida de una cantidad de trabajo cuyo crecimiento es lineal. Hace falta cierto número de segundos para determinar la posición de cada clavo y para clavarlo en el tablero. Una vez la banda ajustada y en su sitio, la operación de anotar sus vértices requiere tiempos algo inferiores al lineal, pues tan sólo hay que registrar las posiciones de los clavos tocados por la cinta elástica.

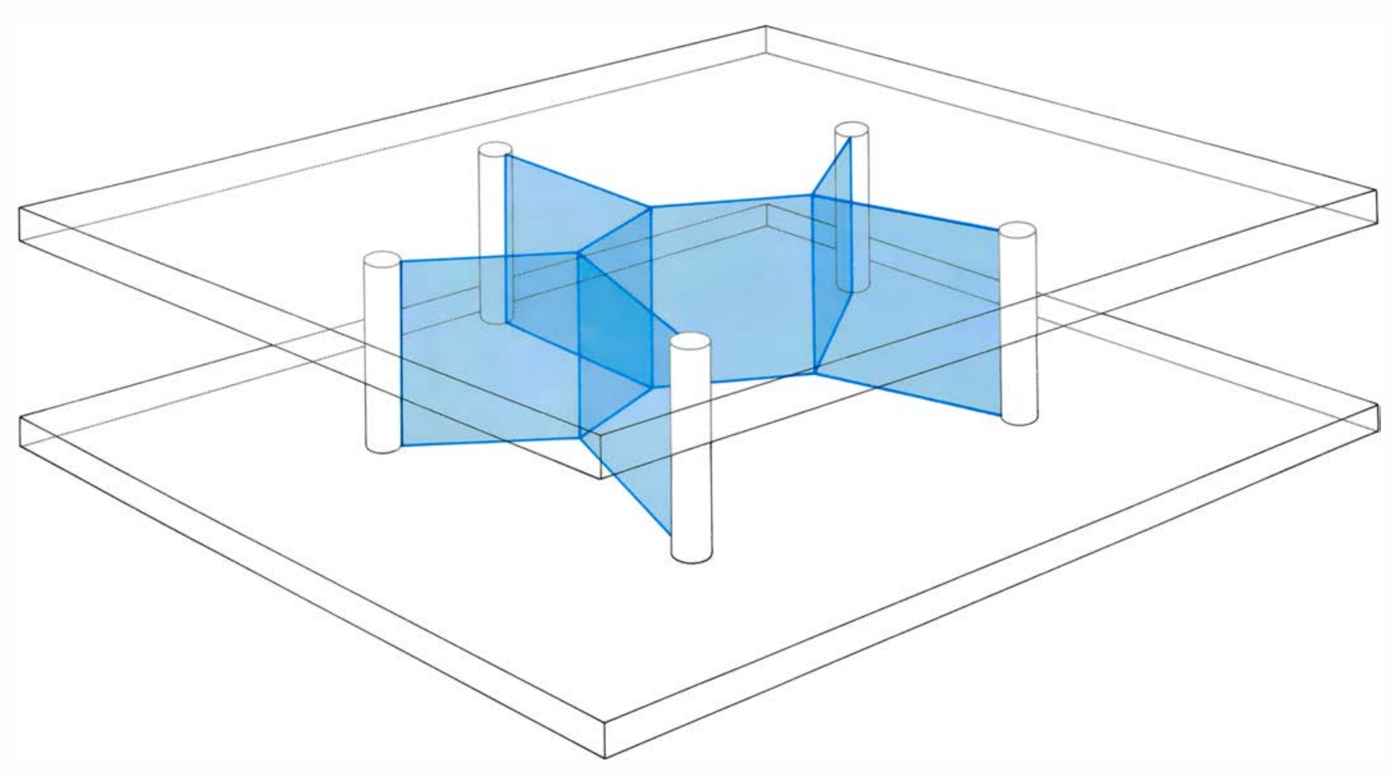
El más veloz de los algoritmos digitales que para hallar la envoltura convexa se conoce requiere del orden de $n \log (\log n)$ pasos. Tan cercano está este valor del lineal, que una CABE capaz de competir con un ordenador digital que ejecutase el algoritmo tendría que ser del tamaño del sistema solar.

La función de un tercer ingenio analógico consiste en hallar el camino más corto entre los que conectan pares de vértices de un grafo. En este contexto, llamamos grafo a una red de líneas –conocidas también por lados– que unen los puntos denominados vértices. En algunos grafos los lados pueden tener longitudes distintas; sobre grafos de este tipo opera el DAC (Dispositivo Analógico a Cordel), que se vale de cordeles finos para materializar grafos. Específicamente, cada vértice está representado por una diminuta argollita de latón. Cuando dos vértices están unidos por un lado del grafo, las correspondientes anillas se conectan por una pieza de cordel, cortado a la longitud correcta y provisto de un ganchito en cada extremo.

Para hallar el camino mínimo entre dos vértices se alza la malla por las ani-



4. Artilugio analógico para hallar el camino más corto entre dos vértices de un grafo



5. *Una solución del problema del árbol de Steiner mediante películas jabonosas*

llas u y v, sujetando una en cada mano. Seguidamente, se tensa desde ellas la red. Instantáneamente destaca en lo alto de la red el camino más corto, formado por una sucesión de hilos tensos: todos los hilos correspondientes a caminos que no sean el mínimo permanecerán más o menos flojos [*véase la figura 4*]. Rotulando las anillas con los nombres de los vértices que representan podremos leer sin dificultad los hitos que definen el camino mínimo.

Si medimos el tamaño del grafo por el número de vértices que lo componen, resulta que las fases preoperatoria y postoperatoria requieren a lo sumo un número linealmente creciente de pasos. La fase analógica del cómputo es virtualmente instantánea. El más rápido de los algoritmos digitales conocidos para este problema exige del orden de n^2 pasos.

La inclusión de las fases previa y posterior al cómputo analógico resulta levemente irritante al analizar las velocidades relativas de estos artilugios, pero por honradez me veo obligado a incluirlas en la descripción del procedimiento analógico completo. ¿Cómo, si no, especificar un problema, o interpretar los resultados obtenidos con el artilugio?

Hay dispositivos que para nada requieren procesamiento, porque están integrados en sistemas que utilizan directamente los resultados. Dos ejemplos que se me vienen a la mente son el clasificador de aguja, antes muy utilizado en las bibliotecas, y un truco para sostener en equilibrio una bandeja sobre la punta de un dedo. En el clasificador de aguja se utiliza un juego de fichas especiales, en uno de cuyos costados se ha practicado una serie de muescas y agujeros, con el fin de saber qué libros habrán de devolverse en una fecha dada. Se apilan las tarjetas, formando un rimero, y por el agujero correspondiente al día se inserta una larga aguja. Al alzarla, tan sólo arrastrará consigo las fichas que en esa posición tengan un agujero y no una escotadura; estas últimas dejan escapar la aguja y permanecen en el mazo.

El truco del equilibrio me lo enseñó Ronald L. Graham, de AT&T Bell Laboratories. Graham comienza por sostener en lo alto la bandeja, apoyándola en el pulgar y en otros dos dedos, muy separados. Conforme los va acercando, el centro de gravedad permanece dentro del triángulo definido por ellos, porque los dedos se deslizan tanto más fácilmente cuanto menos peso soportan. La intrusión final de otro dedo permite acertar lo suficientemente bien con el centro de gravedad como para sostener sobre él la bandeja en equilibrio.

Hasta ahora he venido aplicando estos artilugios a problemas que disponen ya de una solución algorítmica razonablemente rápida. Se dice que tales problemas son de "complejidad polinómica", porque el número de pasos necesarios para resolver un caso particular de tamaño n puede expresarse (o al menos, acotarse superiormente) por medio de una función polinómica de n. Por ejemplo, el tiempo preciso para la solución puede ser proporcional a n, o a n^2, o a $n \log n$, o a n^{27}. Hay otros problemas, entre ellos algunos de importancia práctica, para los que no se sabe de algoritmos de "tiempo polinómico"; los mejores algoritmos conocidos parecen exigir una duración de proceso que crece en función exponencial del tamaño del problema, típicamente como 2^n, función cuyo crecimiento es más rápido que el de cualquier polinomio. El tiempo necesario para casos particulares de gran tamaño de problemas de este tipo, e incluso en los de tamaño moderado, es exorbitante. ¿Por qué no intentar resolver estos difíciles problemas mediante algún artilugio?

Un problema que podría servir de candidato es hallar el camino más largo que conecta dos puntos de un grafo, problema que resulta mucho más difícil que el de hallar el camino mínimo. Se dice, en efecto, que el problema es *NP*–completo, propiedad que al parecer lo condena a la eterna intratabilidad por métodos algorítmicos.

En ocasiones, un poco de fuerza bruta puede hacer entrar en razón a uno de estos difíciles problemas. Tomemos la red de cordeles que antes tensa-

mos para resolver el problema del camino más corto. Si tiramos con más fuerza todavía llegará a romperse alguno de los cordeles, y luego, otro y otro. Al cabo, justamente antes de que la malla quede desgarrada en dos, todos los cordeles que aún restan intactos permanecerán en tensión. Acabamos de resolver el problema del camino de longitud máxima.

¿Seguro? El método resulta eficaz para ciertas redes, pero no para otras. Los lectores encontrarán amena la búsqueda de ejemplos de grafos de uno y otro tipo. Cada vez que al estirar quede en tensión más de un hilo, debe suponerse que se rompe el menos favorable.

Tal vez deberíamos tener más seso, y no intentar dar con elegantes y espectaculares soluciones analógicas a problemas conocidos como *NP*-completos. Mas la teoría de computabilidad lo único que dice es que los problemas *NP*-completos les resultan duros de roer a los ordenadores digitales; nada hay que indique la imposibilidad de hacerlos ceder por métodos analógicos. Por consiguiente, volvamos a intentarlo.

El problema del árbol mínimo de Steiner pide conectar n puntos del plano por medio de un grafo de longitud total mínima. El problema es sustancialmente el mismo que comunicar mediante una red de carreteras n ciudades, con trazado total de longitud mínima. Es lícito tomar como vértices del grafo no sólo los puntos dados, sino también otros que puedan convenir, que, por así decirlo, equivaldrían a nudos y cruces de las carreteras en campo abierto. Es fácil ver que el grafo mínimo tiene que ser un árbol, esto es, un grafo desprovisto de bucles cerrados. Resulta, además, que cualesquiera vértices adicionales del árbol mínimo estarán unidos a otros tres vértices, mediante lados que formarán entre sí ángulos de 120 grados [*véase la figura 5*].

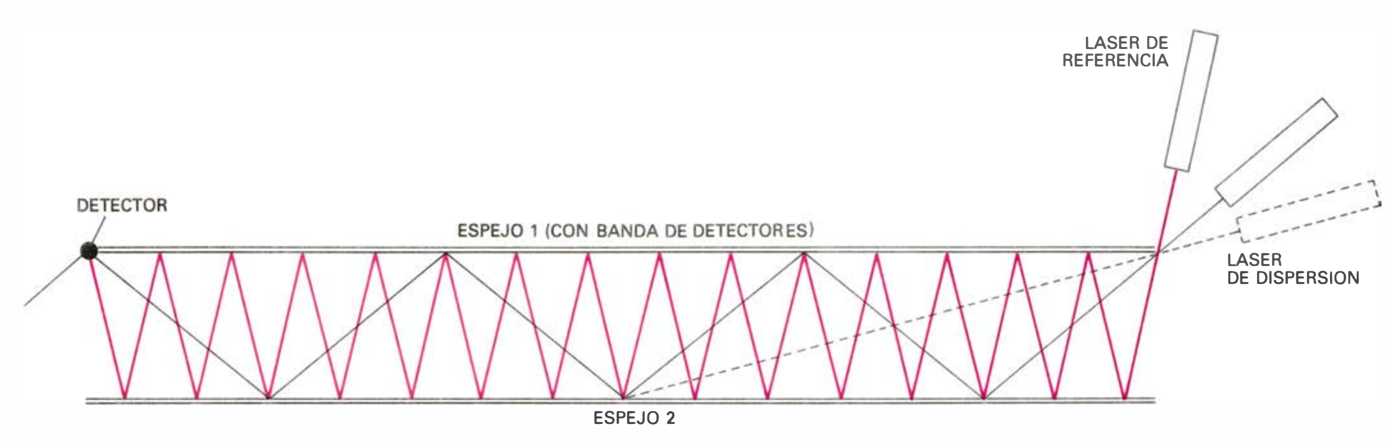
El problema del mínimo árbol de Steiner es *NP*-completo, y no existen para él algoritmos conocidos que exijan menos de 2^n pasos computacionales (siendo n el número de vértices). Hay, empero, un curioso dispositivo, que llamaremos DIAPO (Dispositivo Analógico de Pompas de jabón) que da al parecer fluidas soluciones al problema. Adaptemos un asa a un par de láminas paralelas de plástico rígido y transparente, e insertemos entre las láminas alfileres que representen los puntos a interconectar. Ahora bañemos el artefacto en una solución jabonosa, y saquémoslo. Se forma una película jabonosa que conecta los n alfileres, formando una preciosa red arbórea de Steiner.

Empero, el gozo que produce la anterior solución es prematuro. ¿Cómo estar seguro de que el árbol generado es de longitud mínima? Acontece que la película jabonosa crea siempre la red de longitud total mínima para una topología dada, pero tal vez exista otra topología en la que pueda formarse un árbol todavía más corto. Según la configuración de los alfileres y el ángulo con que el artilugio se extrae de la disolución, la longitud de la red pelicular puede, o no, ser la mínima absoluta. Nuevamente, el intento de resolver un problema *NP*-completo por métodos analógicos se ha quedado en nada.

Uno de los problemas computacionales más famosos es el de establecer si un número entero dado, N, es primo o compuesto. (Se dice que un número es compuesto cuando admite, además de 1 y de sí mismo, otros divisores; en caso contrario, es primo.) Hay una solución analógica muy rápida a este problema, que utiliza el Calculador Analógico Lasérico (CAL). Situamos dos espejos paralelos, *E1* y *E2*, y dos láseres, de modo tal que ambos reflejen su luz muchas veces entre *E1* y *E2*, como muestra la figura 6. El ángulo de uno de los láseres se ajusta de modo que su luz incida N veces sobre cada espejo, y finalmente, ataque un detector situado en el extremo de *E1*. El segundo láser se ajusta de modo que su haz luminoso coincida inicialmente con el del primero, aunque luego se le hará girar un cierto ángulo.

Conforme va variando el ángulo del segundo haz, va periódicamente incidiendo en el detector situado al extremo de *E1*. El espejo *E1* va provisto también de una banda detectora que corre a todo lo largo de él, con lo cual, cada vez que los dos haces coinciden en el mismo punto, la coincidencia es detectable. Un circuito electrónico sencillo se ocupa de supervisar al mismo tiempo el detector situado al final del espejo y la banda detectora. Cuando ambos dan cuenta de una coincidencia simultánea, el circuito hace que se encienda una señal luminosa. Tal señal indica que el número no es primo, sino compuesto; pues para que haya podido producirse la doble coincidencia, el segundo láser habrá tenido que hacer reflejarse su luz n veces en el espejo *E1*, y n tendrá que dividir exactamente a N.

Al hablar de la dificultad de decidir si un número N es primo o compuesto hay que ser cuidadoso en la medición del tamaño del problema. Definiéndolo simplemente como N, existen algoritmos digitales que permiten resolverlo en tiempos polinómicos. Se considera, empero, más correcto definir el tamaño como log N, porque ésta es la longitud de la cadena de dígitos necesarios para la representación de N. Con esta medida se ignora si el problema de la determinación del carácter primo o compuesto de un número es un problema *NP*-completo, o no. Basándonos en el éxito de la computadora CAL, y en el fracaso de otros artilugios analógicos para resolver problemas digitalmente intratables, ¿podríamos conjeturar que el problema no es *NP*-completo?



6. Prueba el Calculador Analógico Lasérico que 15 no es número primo

Taller y laboratorio

Chismes que aplican principios físicos no inmediatos al placer de cocinar

Jearl Walker

La base del arte culinario no es otra cosa que física (química incluida), aunque casi todo el mundo (yo, también) cocina apoyándose en un instinto guiado por la experiencia. No obstante, cuando hay posibilidad de echar mano de algún principio de la física que facilite o perfeccione un guiso la cosa resulta más divertida. Hace poco dí con varios artefactos de cocina cuyo funcionamiento se basa en ideas no evidentes de suyo. Entre ellos se cuentan una pipa de calor que acelera el guisado de la carne, un cristal líquido que controla la cocción de los huevos, una mantequera capaz de mantener fría la mantequilla en un lugar cálido y una olla china que cocina al vapor aunque parte del alimento sea líquido.

Una pipa de calor para cocina consiste en un tubo hueco cerrado, normalmente con un extremo aguzado para que pueda atravesar la carne. En su interior contiene una pequeña cantidad de un fluido cuya misión es transferir calor desde el horno hacia el interior frío de la carne. Como esa transferencia es más rápida que la velocidad a la que se conduce el calor a través de la carne, se reduce el tiempo de cocción y disminuye el encogimiento que produce toda cocción prolongada.

En mi pipa de calor preferida el extremo exterior está incrustado en un cilindro metálico macizo. Tengo una docena de pipas más, todas ellas más largas y con los cilindros del extremo más pequeños. Las pipas de calor específicas para cocina tienen aletas de refrigeración. Cilindros y aletas cuya misión consiste en absorber calor del horno para que se vaporice el fluido interior de la pipa. El vapor, cuando se esparce por la pipa, se condensa en parte en la zona rodeada por el interior relativamente frío de la carne, cediendo calor a la pared de la pipa y a la carne circundante.

Normalmente, la pipa se coloca en la carne en posición sesgada para que el fluido condensado regrese al fondo por gravedad. Para ayudar a ello la pipa reviste su interior de un material poroso que arrastra al líquido por capilaridad. En el fondo el líquido se vaporiza y el ciclo comienza otra vez. Cuando la parte superior de la pipa se calienta, disminuye la condensación y el ciclo se decelera. Y, cuando el interior de la carne está casi tan caliente como el horno, cesa la circulación por la pipa.

El líquido encerrado en el útil puede ser agua, metanol u otra sustancia que requiera mucho calor al vaporizarse. Consideremos una pipa que contenga un gramo de agua a la temperatura ambiente (19 grados Celsius, por ejemplo). Para calentar ese agua hasta la temperatura de vaporización normal de 100 grados Celsius se necesita una energía de 340 joule, lo que no es demasiado en comparación con el calor necesario para vaporizarla (2256 joule). Toda esa energía se necesita para liberar las moléculas y que formen un gas.

La energía necesaria para vaporizar el agua suele llamarse calor latente de vaporización; es la que transporta la pipa de calor para acelerar la cocción de la carne. Cada vez que el agua se condensa en una zona fría del utensilio pierde ese calor latente cuando las fuerzas que se ejercen las moléculas vuelven a formar el líquido. Entonces, el calor liberado por el agua atraviesa la pared de la pipa con destino a la carne circundante. O sea, si cada segundo hay un gramo de agua que deposite su calor latente, la carne recibirá 2256 joule por segundo.

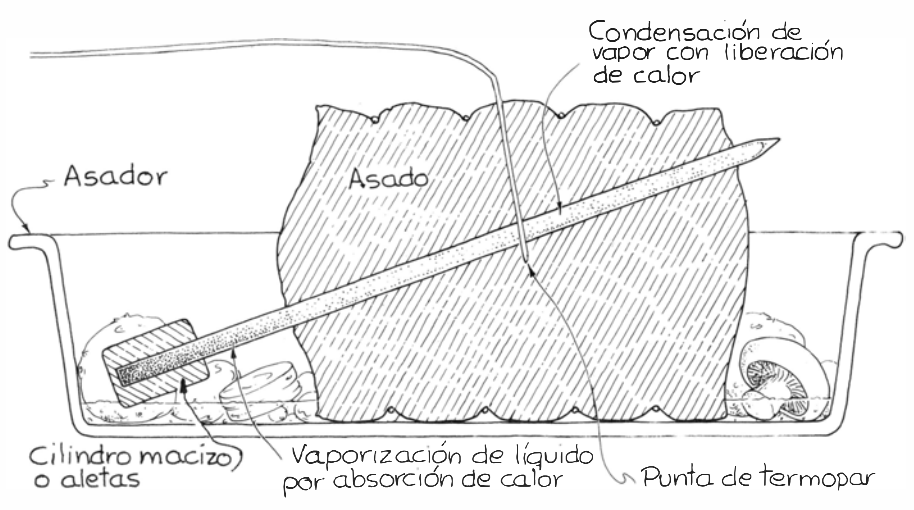
A título comparativo hice el cálculo de la cantidad de calor que conduciría hasta la carne una barra de aluminio aleado maciza. Si ésta tuviera el mismo diámetro que mi pipa de calor depositaría únicamente 35 joule por segundo, bastante menos que la pipa.

Efectué dos ensayos. Una noche asé un pequeño solomillo (de un kilo de peso aproximadamente). Recurrí a mi vieja pipa, la del cilindro macizo en el extremo, con la que traspasé la carne de un lado a otro y hacia arriba, hasta que asomó la punta afilada.

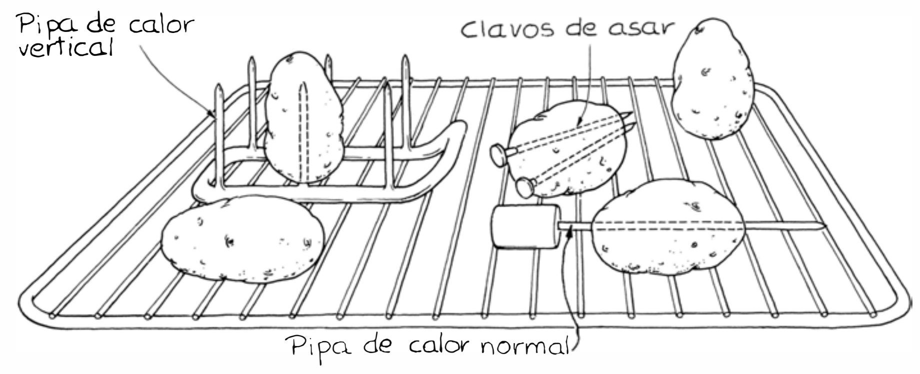
Con un clavo practiqué un agujero en la parte superior de la carne para instalar en él un termopar detector de calor. Su profundidad era de unos 3,5 centímetros y ceñía ajustadamente los cables del termopar. Así podía medir la temperatura a unos dos centímetros de la pipa, que en ese punto se encontraba a 2,5 centímetros de la superficie superior del asado, y aproximadamente en el centro respecto a su longitud y anchura. La temperatura del termopar la registraba en un termómetro de termopar.

Coloqué el asado en la bandeja central de mi horno eléctrico, que había precalentado a un poco más de 200 grados Celsius. En los primeros 25 minutos la temperatura medida en la punta del termopar se incrementó en 26 grados, para alcanzar un aumento total de 58 grados en 44,5 minutos.

En ese instante retiré el asado del horno y lo corté en varias tajadas verticales. El color de la carne me dio una pista acerca del grado de cocción. La tajada central estaba tostada sólo algunos milímetros por debajo de la parte superior de la pieza y dos centímetros por encima de la parte inferior. El



1. Pipa de calor para acelerar la cocción de la carne



2. Experiencias en el asado de patatas

agujero que había dejado la pipa aparecía rodeado de una zona tostada de unos dos centímetros de diámetro. Pero el resto de la tajada seguía roja. (Debo advertirles que tan costoso material de laboratorio no se desperdició. No era caso, obviamente, de servirlo en una mesa, pero encontré el modo de convertirlo en algunos manjares exquisitos que no hace falta describir ahora.)

Una tajada vecina al punto de entrada de la pipa estaba más tostada en torno al agujero que había hecho el utensilio. Y otra, situada al otro extremo del trozo de carne, lo estaba menos. Por tanto, la transmisión de calor a la carne es más eficaz cerca del extremo inferior de la pipa que en el extremo superior de la misma, posiblemente porque el líquido que se condensa en el primero ha de recorrer menos distancia para regresar al fondo y, por ello, regresa antes.

A la noche siguiente asé una pieza similar del mismo peso aproximado, pero sin la pipa. Hundí el termopar más o menos en el mismo sitio; seguí la misma marcha de cocinado. Esta vez, en los primeros 25 minutos, la temperatura del termopar aumentó sólo 12 grados y el incremento de 58 grados tardó una hora y nueve minutos, del orden del 50 por ciento más que con la pipa. (Esta vez nos comimos el asado. Dicho sea de paso, si el lector se pregunta si una pipa de calor favorece más la cocción de la masa de una pieza de carne, le contestaré que sí. Esto puede sonar terrible para un asado de buey, pero es una clara ventaja para asar cerdo. Sobre gustos no hay disputa.)

Existen en el mercado asadoras de patatas. La que yo compré está provista de seis pipas de calor verticales de unos nueve centímetros de largo. Se ensartan allí las patatas y se introduce la bandeja en el horno. Como las patatas conducen muy mal el calor, muchas veces les pongo clavos, confiado en que al conducir el metal el calor con mayor rapidez, se acelerará la cocción. Por ello, creí que una pipa de calor transmitiría el calor más deprisa todavía.

Para verificar estas suposiciones hice un experimento con cinco patatas del mismo tamaño (11×6×5 centímetros). Las coloqué todas muy juntas en la bandeja central de un horno precalentado a un poco más de 200 grados Celsius. El costado de cada una lo perforé con un clavo: el termopar recogería la temperatura a una profundidad de dos centímetros. Puse una patata horizontal, vertical otra; le inserté a una tercera dos clavos de aluminio especiales para asar patatas y ensarté otra en una pipa de calor vertical. Traspasé la quinta con mi antigua pipa de calor y la coloqué horizontalmente sobre la bandeja. (Como la pipa resultaba muy larga para la patata, desgarró un poco uno de sus extremos.)

Transcurridos 10 minutos comprobé la temperatura de cada patata. La de la pipa grande se había calentado 28 grados, la horizontal 23 grados, la instalada en la pipa vertical 19 grados, y, 17 grados, las otras dos. Veinte minutos después, la patata de la pipa grande seguía en cabeza, pero las demás presentaban unas temperaturas muy similares. La patata vertical permanecía la más fría.

Proseguí con la carrera hasta que las patatas llegaron aproximadamente a los 100 grados Celsius. Con sus 55 minutos, ganó la de la pipa grande. Las otras cuatro tardaron unos 74 minutos en quedar listas.

Los clavos para asar no rebajaron el tiempo de cocción. Si bien conducen rápidamente el calor, no introducen mucho en el vegetal. Aunque éste se asa antes con una pipa de calor, no conviene que se ponga vertical, pues así presenta una sección transversal reducida a las bobinas calefactoras del fondo del horno e intercepta una fracción pequeña de la radiación infrarroja directa. El calor que transporta la pipa puede quedar neutralizado por esta pérdida de absorción de radiación directa. Lo mejor es colocar las patatas horizontalmente sobre las bobinas. Estas conclusiones mías son provisionales. Si alguien prueba con estos ensayos, recibiré con gusto noticias de sus resultados. (Nosotros sacrificamos las patatas.)

Hasta hace poco, el tiempo de cocción de los huevos lo calculaba a ojo. Jamás podía estar seguro del estado que habían alcanzado. Pero un nuevo utensilio ha venido a resolver limpiamente el problema. Se trata de un ingenioso cronómetro para huevos. Está hecho de plástico transparente. En el centro lleva una delgada película roja con una escala en que se lee "pasado por agua", "medio hecho" y "duro".

El cronómetro se mete en el agua justo con los huevos. Cuando sube la temperatura, la película roja comienza

Avance del cambio de calor
Penetración de calor en un huevo
Agua hirviendo
DURO
MEDIO
PASADO POR AGUA
EGGRITE

3. Cronómetro para huevos

a cambiar de color desde el rojo vivo al rojo oscuro. Esta variación de color empieza notándose en el borde de la película y, conforme pasa más calor a través del plástico, se va oscureciendo más superficie de película. El cambio de color no tarda en llegar a la señal de "pasado por agua", indicando que tal es el estado de los huevos. Calentando más, el cambio de color avanza hacia el centro de la película, pasando por "medio hecho" y "duro". Con un poco de experimentación se puede calibrar la escala para conseguir huevos de cualquier grado de consistencia. También se puede calibrar la escala en función del tamaño de los huevos.

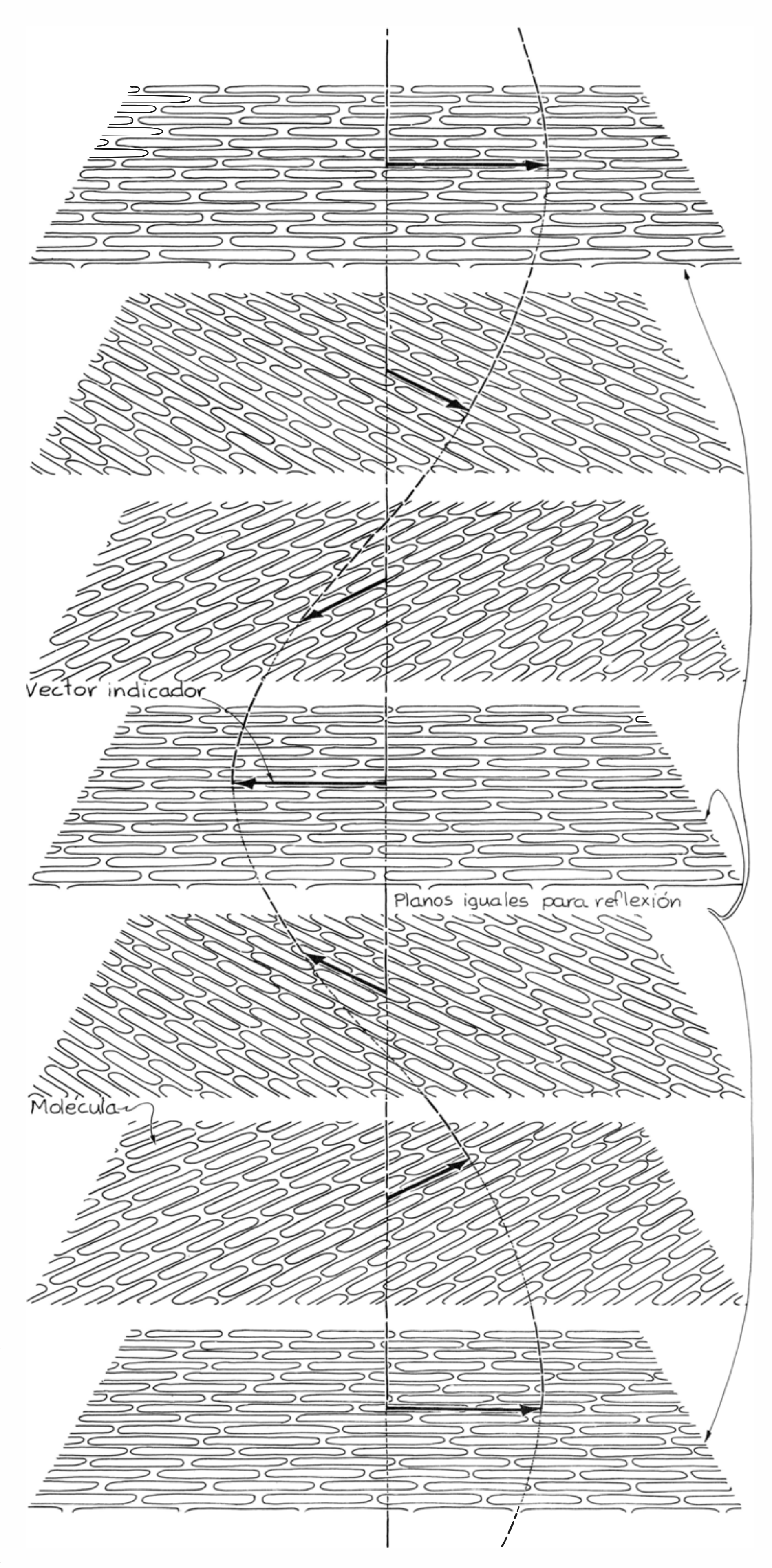
El cronómetro tiene la misma sensibilidad que los huevos a la temperatura inicial del agua, al número y tamaño de éstos que hay en la cacerola y a la velocidad de calentamiento del agua. Suponiendo que rebajemos ésta reduciendo el calor o añadiendo huevos, la conducción de calor hacia éstos se hace más lenta y también la variación de color.

El instrumento responde incluso a los cambios de presión atmosférica. Supongamos que lo llevemos desde el nivel del mar hasta una altitud considerable. Como la presión atmosférica disminuye con la altitud, la temperatura de ebullición del agua baja. Entonces, cada huevo recibirá menos calor por segundo y, por tanto, tardará más en cocerse. Pues lo mismo le ocurre al cronómetro.

Para estudiar la película abrí el aparato cascándolo con un pequeño martillo. (Con la precaución de protegerme los ojos de los pedazos que salieran proyectados.) La película consta de dos capas. La de arriba es de plástico transparente flexible. La de abajo la forma una sustancia roja esponjosa que parece haber sido pintada sobre la de arriba. Valiéndome de unas tenazas sumergí la película en agua y calenté esta, a la vez que medía su temperatura con un termopar.

Cuando el agua vecina a la película alcanzó unos 68 grados Celsius, ésta viró rápidamente del rojo vivo al rojo oscuro, manteniendo esta tonalidad hasta el punto de ebullición del agua. Entonces alcé la película del agua y la toqué en un borde. El efecto refrigerante de mi dedo hizo que el color de éste cambiará inmediatamente a rojo vivo. Con agua fría retornaba el rojo en toda su viveza.

Yo diría que la capa esponjosa inferior consiste en un cristal líquido, probablemente de tipo colestérico. En estas sustancias las moléculas guardan cierto orden cristalino, aunque la sustancia sea fluida. Mi opinión es que el



4. Disposición en capas de las moléculas de un cristal líquido colestérico

cristal líquido del cronómetro refleja de manera intensa la luz roja hasta los 68 grados Celsius aproximadamente. Luego, para temperaturas más altas, absorbe la mayor parte de la luz y se torna rojo oscuro.

La estructura de los cristales líquidos colestéricos es estratificada y las moléculas, a modo de bastoncitos, yacen en cada capa con los ejes mayores en los planos de éstas. La orientación de las moléculas varía gradualmente de capa en capa. Así, si tomamos un vector que apunte en la dirección de los ejes mayores de las moléculas de cada capa, este vector describirá un trayecto helicoidal al desplazarse a través de las sucesivas capas.

El cristal podemos imaginarlo compuesto de planos espaciados regularmente que reflejan la luz ambiente. La separación entre planos es la distancia que debe ascender el vector indicador a través de las capas para girar 180 grados. Como se considera que las moléculas son bastoncitos con extremos indistinguibles, esos dos planos presentan moléculas de igual orientación.

Supongamos que un conjunto como ese de planos regularmente espaciados sea iluminado por luz blanca. La luz reflejada en una capa interferirá la reflejada en otra. Para la mayoría de las longitudes de onda, aquellos rayos reflejados se interfieren destructivamente para dar oscuridad, o como mucho, un calor pálido. Para otras longitudes de onda, los rayos se interfieren constructivamente y producen un color vivo correspondiente a esas longitudes de onda.

Es frecuente que los cristales líquidos colestéricos sean sensiblemente fluidos e iridiscentes. Los colores de las distintas partes del fluido dependen de la orientación de los planos del cristal en dichas partes. Al calentar una muestra varía la distancia entre los planos y, con ello, el color que se devuelve con más intensidad al observador. Si la temperatura es suficientemente elevada, las vibraciones moleculares desorganizan la estructura del cristal y el fluido pierde su aptitud para reflejar intensamente ciertos colores.

En lo referente al cronómetro yo afirmaría que se han mezclado otras sustancias con un cristal líquido y, después, se ha pintado sobre la capa de plástico flexible. Esa mezcla, que no es iridiscente, al iluminarla con luz blanca devuelve una fuerte luz roja al observador. A unos 68 grados Celsius se pierde tal retorno selectivo de luz roja. Ahora se absorbe la mayoría de la luz incidente y la película queda roja oscura. En cuanto se enfríe ésta, se restablecen la estructura ordenada y la reflexión selectiva.

La mantequillera de terracota es un recipiente de arcilla procedente de Italia que retrasa la tendencia de la mantequilla a fundirse en los días cálidos. Yo la empleo en mis meriendas en el jardín de casa. Sin ella, la mantequilla se ablanda enseguida y con ella permanece dura durante casi toda la fiesta.

El plato inferior del recipiente está vidriado para que la mantequilla no impregne la arcilla. De la regulación de la temperatura responde la tapa, que no está vidriada. Alrededor de una media hora antes de comer le doy la vuelta a la tapa y la lleno de agua fría. Hacia la hora de comer, la arcilla está saturada. Entonces extraigo el agua sobrante y coloco la tapa sobre la mantequilla (fresca de la nevera) y el plato de fondo.

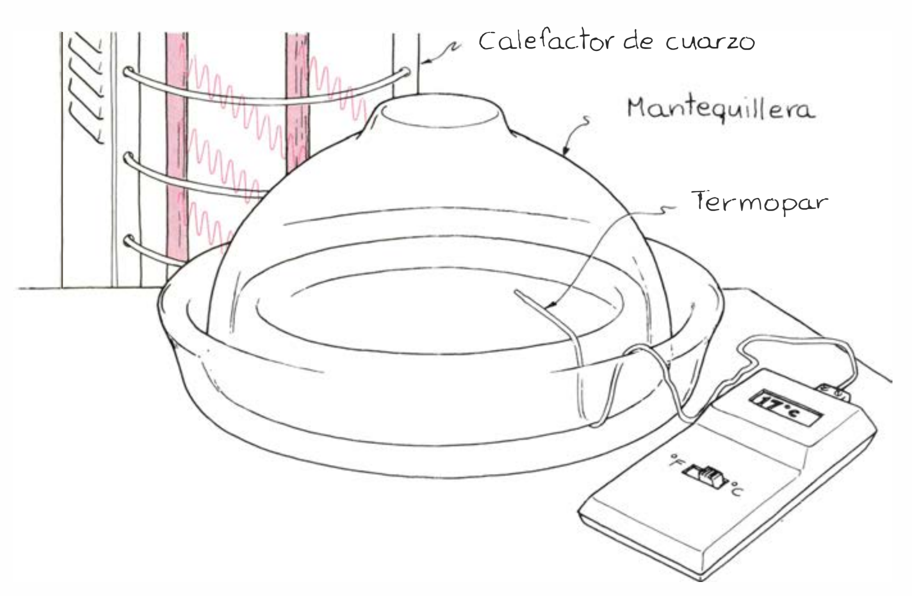
En cuanto el recipiente se lleva al aire libre comienza a calentarse, pues absorbe calor del entorno. Aunque no esté expuesto a la luz solar, todavía absorberá radiación visible e infrarroja. También recibirá calor por convección del aire que se mueva en torno suyo. Y en el recipiente penetrará aún más calor procedente de la mesa, a menos que se coloque sobre un aislador, un salvamanteles por ejemplo.

Si el recipiente está completamente seco, la energía tomada de la radiación y del aire exterior se transmite gradualmente al interior a través de las paredes. Cuando la tapa está húmeda, la conducción de calor por la pared se hace más lenta: gran parte de la energía se consume en la evaporación del agua. A medida que el exterior se seca, del interior de la pared sale más agua para completar la existente en el exterior. Así, con el retraso en la transmisión del calor, permanecen fríos el aire interior y la mantequilla.

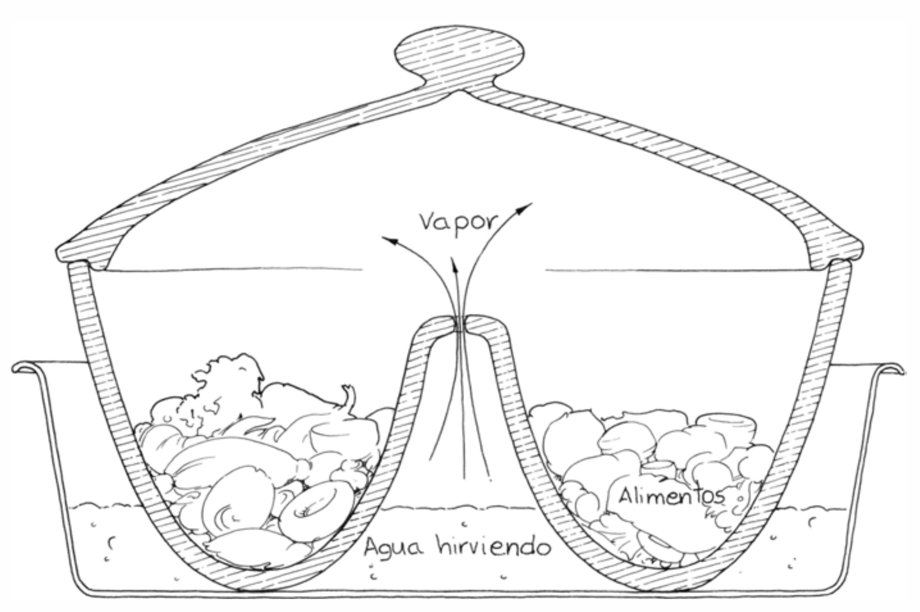
Para contrastar esta explicación preparé un experimento en el que calentaba la mantequillera con un calefactor de cuarzo, que emite radiación infrarroja y algo de radiación visible. Primero empapé la tapa de arcilla dándole la vuelta y llenándola de agua. Su superficie exterior aparecía evidentemente seca. Pasados dos o tres minutos, el agua que llenaba la tapa generó un gran número de burbujas, conforme el agua expulsaba el aire de la arcilla. Al cabo de 45 minutos el agua había empapado toda la pared y humedecía la superficie exterior.

Tras vaciar la tapa, la instalé en el plato e introduje un cable termoeléctrico en el interior del recipiente. La punta del termopar se encontraba junto al centro del plato y a pocos milímetros por encima. Con él podía registrar la temperatura del aire interior del recipiente. Este y el termopar los coloqué entonces a unos 30 centímetros delante del calefactor de cuarzo, el cual proporcionaba más calor del que cabría esperar en un día de verano. Así, por un costado, la mantequillera estaba bañada de radiaciones infrarroja y visible; por el otro se encontraba ante una fría habitación de sótano, de la que recibía una radiación infrarroja relativamente escasa.

Al recibir calor el recipiente, la temperatura de su aire interior comenzó a subir. Inicialmente a razón de un grado centígrado cada cinco minutos. Media hora más tarde esa velocidad bajó hasta un grado cada ocho minutos. Para entonces la temperatura del aire había aumentado únicamente cinco grados y el costado enfrentado al calefactor aparecía seco a la vista y al tacto. Por su



5. La mantequillera de terracota



6. Olla Yunnan para cocinar al vapor

parte, el costado oculto al calefactor aparecía aún húmedo.

Dejé que el recipiente se secara durante un día y repetí la experiencia con la tapa seca. Esta vez el aire interior se calentó a razón de un grado cada tres minutos. Media hora más tarde, la temperatura se había elevado 10 grados, el doble que con la tapa húmeda.

Luego pesé la tapa antes y después de empaparla con agua. La tapa empapada acumula del orden de 85 gramos de agua (tras vaciarla, por supuesto). Para evaporar esa cantidad de agua hay que aportar una energía de 192 kilojoule aproximadamente. Admitiendo que la intensidad de la luz solar sobre la mesa donde organizo mis meriendas sea del orden de un kilojoule por segundo y metro cuadrado, la tapa debe tardar unas dos horas en captar sol suficiente para secarse del todo. La verdad es que se seca antes, porque los objetos próximos producen también radiación infrarroja y visible, y a causa del calor que ceden la convección y la conducción.

En la olla de la clase Yunnan se emplea también agua, pero con distinto objeto. El utensilio es una olla de arcilla vidriada provista de una chimenea central por cuyo interior sube vapor que calienta el alimento encerrado en ella. Para cocinar carne o verduras picadas, instalo la marmita sobre una cazuela honda llena de agua. Cuando el agua se calienta, por la chimenea sube el vapor que cuece poco a poco las viandas.

La transferencia de calor es casi igual que en la pipa de calor; con una salvedad: el agua no regresa al foco calorífico, sino que se reúne en la olla con el alimento y lo mantiene humedecido. Es más común cocinar al vapor con un cesto metálico perforado que se instala directamente encima de agua hirviendo. En este caso, el vapor se eleva atravesando los orificios del cesto y se condensa en el fondo de la comida, cediendo a ésta su calor latente de vaporización.

Parte del agua condensada se adhiere a los alimentos, pero la mayoría se escurre hasta el fondo de la bandeja y vuelve a vaporizarse. Entonces, la comida se humedece pero no se empapa. La ventaja de la olla Yunnan sobre el cesto metálico es que permite cocinar al vapor un alimento que esté ya líquido. Con un cesto el líquido se perdería en la masa de agua.

Para examinar la marmita ideé dos sencillas experiencias, en las que utilicé una sonda termopárica que instalé en su interior, a unos tres centímetros del fondo y en un punto aproximadamente equidistante de la pared y la chimenea. Inicialmente la temperatura del aire interior a la olla era de 21 grados Celsius. Cuando la introduje en una cazuela de agua hirviendo, la temperatura del aire empezó a subir a razón de 0,6 grados por segundo, llegando a los 90 grados tras poco más de tres minutos.

Después de enfriarla tapé el extremo superior de la chimenea con cinta adhesiva gruesa y repetí las medidas. Entonces el incremento de temperatura resultó unas cuatro veces más lento. Pasados tres minutos, la temperatura no alcanzaba todavía los 55 grados. Evidentemente, el aumento rápido de temperatura dentro de la olla, normal, lo fuerza el calor latente de vaporización del vapor de la chimenea. La conducción del calor por el fondo del recipiente afecta menos.

Libros

Medicina y biología en busca de un estatuto epistemológico

Pedro Laín Entralgo y Guillermo Olagüe de Ros

INTRODUCCIÓ A L'ESTUDI DE LA MEDICINA EXPERIMENTAL, por Claude Bernard. Traducció i Pròleg de Jaume Pi Sunyer. Biblioteca Clàssics de la Ciència. Edicions Científiques Catalanes, Barcelona, 1983. Puesta brevemente al día por un prólogo del mismo traductor, compuesto "en els anys plàcids del capvespre", la excelente traducción catalana de este inmortal libro que en 1936 publicó la casa editorial Arnau de Vilanova ha sido reeditada, bajo el patrocinio de la Generalitat de Catalunya, por Edicions Científiques Catalanes. Buena ocasión, pues, para un rápido examen de lo que la *Introduction à l'étude de la médecine expérimentale* dice al lector actual.

Nadie ha ensalzado tanto el valor de esta *Introduction* como el filósofo H. Bergson: "A estas investigaciones concretas (las de las ciencias de laboratorio) –escribió– aportará Cl. Bernard la fórmula de su método, como antaño Descartes a las ciencias abstractas de la materia. En este sentido, la *Introducción al estudio de la medicina experimental* es para nosotros un poco lo que para los siglos XVII y XVIII fue el *Discurso del método*." Las ideas rectoras de aquélla –el determinismo como presupuesto mental del hombre de ciencia, la introducción de la ley física como meta del conocimiento, el valor de la "idea *a priori*" como punto de partida del método experimental, la validez del método comparativo en la investigación biológica– serían desde entonces el soporte conceptual de todo saber científico basado sobre el experimento. ¿Es así?

En cierto sentido, no. Examinemos cinco de las ideas básicas del genial fisiólogo:

1.ª En microfísica, el riguroso determinismo bernardiano –esta doble convicción: que las condiciones de existencia de todo fenómeno natural están determinadas de manera absoluta, y que el análisis experimental puede conocerlas rigurosamente– ha sido sustituido por el indeterminismo a que debe atenerse el "experimento conceptual" de la mecánica cuántica.

2.ª La ley física, que en último término, si a tanto se llega, debe ser cuantitativa y matemática, no es la única meta de la investigación científica; las teorías de la ciencia actual no están únicamente constituidas por "leyes"; también lo están por "modelos". La doble hélice del ADN y la estructura molecular de la membrana de la célula son un buen ejemplo de ello.

3.ª Para el actual teórico de la ciencia, el valor gnoseológico de los resultados del experimento científico más irreprochable no es el que Cl. Bernard les atribuía. Baste pensar en los varios movimientos intelectuales que reciben los nombres de "convencionalismo" (W. James, J. Dewey, E. Mach, H. Poincaré, P. Duhem), "verificacionismo" (M. Schlick, E. A. Singer, A. J. Ayer), "falsacionismo" (K. Popper), así como la crítica de éste (Th. S. Kuhn, I. Lakatos, P. K. Feyerabend).

4.ª El investigador científico actual no comparte la renuente actitud bernardiana ante el empleo de la estadística; no sólo porque en el campo de la microfísica todo conocimiento científico tenga que ser meramente probabilístico, también porque el método estadístico puede ser, en biología, un recurso perfectamente válido para conocer científicamente la realidad.

5.ª La investigación experimental de laboratorio no es y no puede ser, contra lo que Cl. Bernard afirmó, el fundamento último del saber médico. Por grande que haya sido y siga siendo la importancia de la investigación experimental, a la postre fisioquímica, en la edificación del saber médico, éste debe *en último extremo* reposar sobre lo que diga al sanador su relación directa con el enfermo. "Yo considero al hospital sólo como el vestíbulo de la medicina científica, como el primer campo de observación en que debe entrar el médico; pero el verdadero santuario de la ciencia médica es el laboratorio", se dice en la *Introducción*. Sé muy bien que muchos eminentes médicos de hoy se hallarían dispuestos a suscribir estas palabras; pero no creo que lo hicieran, si se decidiesen a pensar con cierta detención y rigor suficiente lo que en realidad es, en su integridad, la práctica de la medicina.

Todo lo cual no niega el grande e imperecedero valor de la *Introducción al estudio de la medicina experimental*. Valor no sólo histórico, esto es, por lo que este libro fue en su tiempo, mas también actual, es decir, por lo que todavía sigue siendo, por lo que su lectura puede decir a cuantos con la necesaria responsabilidad intelectual quieran consagrarse a la investigación en el laboratorio.

Dos ejemplos insignes pueden esclarecer esto que afirmo. El primero, la geometría de Euclides. La validez matemática de las geometrías no euclidianas, ¿ha convertido en antigualla inútil los *Elementos* de Euclides y la geometría sobre ellos construida? En modo alguno. La geometría euclidiana es válida como fundamento de la visión geométrica de la realidad cuando se admiten los presupuestos del observador ingenuo, y conforme a esta reserva deben ser leídos y entendidos, si el lector es matemáticamente exigente, los textos en que se la elabora y expone. Segundo ejemplo: los *Principia* de Newton. Hasta los alumnos de enseñanza media saben hoy que la mecánica newtoniana ha sido sustituida por la mecánica relativista, de uno u otro modo aliada a la mecánica cuántica. ¿Quiere esto decir que los principios de la mecánica expuestos en los *Principia* y que la fórmula newtoniana de la gravitación universal son no más que errores científicos, como lo fue la teoría del flogisto de Stahl, y que por consiguiente deben ser relegados al olvido? Qué dislate. La mecánica newtoniana es un caso particular de la mecánica relativista, y así debe ser estudiada y conocida.

Tal es, a mi juicio, la actual situación de la *Introducción* bernardiana. La concepción actual de la ley física y del experimento científico pone un límite, a la vez real y teorético, a la idea del determinismo que preside y da fundamento a ese libro imperecedero; pero

la pauta metódica del experimento que en él se expone –idea *a priori*, razonamiento experimental, experimento, contraprueba– y la ulterior valoración del resultado así obtenido, son perfectamente válidas en el caso de todos los experimentos biológicos cuya ejecución no implique el manejo de magnitudes microfísicas. El empleo de la estadística en la experimentación biológica es, por supuesto, un ineludible recurso técnico; pero esto no excluye la validez y la fecundidad del punto de vista bernardiano y de la regla en que éste se expresa: profundizar hasta donde sea posible el análisis de cada observación individual. La realidad que el médico conoce y trata, una persona enferma, exige el atenimiento a principios y métodos que exceden esencialmente el campo de los que rigen la experimentación en el laboratorio, aunque ésta tenga como objeto la fisiología de los antropoides más próximos a la especie humana; pero el conocimiento científico de la fisiología del organismo y de la psique del hombre exige ineludiblemente dicha experimentación, y gracias a ella ha avanzado tan fabulosamente desde Harvey hasta hoy. En suma: entre la intelección teorética y experimental de la realidad microfísica, por un lado, y el conocimiento formalmente antropológico de las funciones y las vicisitudes del organismo humano, por otro, la *Introducción al estudio de la medicina experimental* sigue conservando un altísimo valor intelectual. Para el biólogo y para el médico no es un libro definitivo –ninguno puede serlo, si el saber humano es su tema–, pero es y seguirá siendo un libro genial e indispensable.

A la luz de la valiosa traducción que de nuevo les ofrece Jaume Pi i Sunyer, así deben verlo –así lo verán, estoy seguro– los médicos y los biólogos catalanes vocacionalmente preocupados por lo que profesionalmente hacen. (P. L. E.)

LA TEORÍA CELULAR. HISTORIA DE UN PARADIGMA, por Agustín Albarracín Teulón. Madrid, Alianza Editorial (Alianza Universidad, n.º 366), 298 páginas, 1983. En el proceso de formulación de una teoría estequiológica universalmente aceptada y aplicable a toda la materia viva, la celular ha sido cronológicamente la última. Expuesta inicialmente hacia la década de los treinta de la centuria pasada por Mathias Jakob Schleiden y Theodor Schwann, y ampliada posteriormente por Jacob Henle, Albert Koelliker, Robert Remak y Rudolph Virchow, entre otros, conocerá su culminación con las investigaciones de Santiago Ramón y Cajal, que con su teoría de la neurona generalizó al sistema nervioso el carácter celular de toda la materia viva. Apenas han sido precisos, pues, cincuenta años para su formulación tal como hoy la entendemos y, sin embargo, a poco de escasas décadas, tras una incuestionable vigencia de media centuria, se han producido ya los primeros síntomas de "crisis" en el conjunto de esta teoría.

En efecto, la nueva serie de datos aportados por la microscopía electrónica, el estudio de los virus y el desarrollo de las técnicas histoquímicas, por sólo citar algunos, están sentando las bases de un nuevo horizonte intelectual que, hoy por hoy, se mueve entre estas tres aporías: seguir considerando a la célula como el verdadero elemento material de la vida; estimar la posibilidad de entidades vivientes infracelulares y, finalmente, concebir al organismo como una estructura supracelular en el que sus células carecen de personalidad propia, siendo la conjunción federada de esas formaciones la última unidad elemental de materia viva. No ha de resultar, pues, extraño que por las características arriba apuntadas, la historiografía publicada hasta la fecha sobre este paradigma central de la biología actual haya sido muy nutrida, aunque de calidad muy variable.

La obra que nos ocupa es el resultado de prácticamente cuatro años de intenso estudio sobre este paradigma de Agustín Albarracín Teulón, profesor titular de historia de la medicina en la Universidad Complutense de Madrid y director del Instituto "Arnau de Vilanova" del CSIC. A lo largo de este cuatrienio, que abarca de 1978 a 1982, Agustín Albarracín ha ido ofreciendo algunos adelantos de su pesquisa, incluso en fechas previas, como son la Ponencia Oficial y una comunicación al V Congreso Nacional de Historia de la Medicina (Madrid, septiembre-octubre de 1977). Tan continua y dilatada dedicación ha permitido al autor de *La Teoría Celular* evitar algunos de los defectos en que, a su juicio, han incurrido otros estudiosos del tema. Así, rechazando buena parte de esos estudios previos, a los que Albarracín imputa un excesivo factualismo y una gran pobreza interpretativa, amén de fragmentarismo, éste ha pretendido ofrecernos una visión global de la génesis y desarrollo de la teoría celular desde un acercamiento internalista, basado en el análisis de la vida y obra de las grandes figuras que contribuyeron a su formulación primitiva y evolución posterior.

De ahí que el enunciado de muchos de los capítulos en que se divide el libro vaya precedido por el nombre de alguna de esas grandes figuras. Pero, además, Albarracín ha querido complicar positivamente su acercamiento insertando la evolución de la doctrina celular en la dialéctica histórica materia-vida, que, a su juicio, constituye uno de los ejes centrales de la problemática biológica. El último elemento novedoso que Albarracín ha introducido para una coherente explicación evolutiva de la teoría celular ha sido, tal como se advierte en el subtítulo de la obra; el esquema de Thomas Kuhn sobre el nacimiento y crisis de los "paradigmas científicos", aplicado en este caso al de la teoría celular.

Nos encontramos, pues, ante una importante novedad en el panorama historiográfico científico español, tanto por el particular modo con que Albarracín ha historiado tan central paradigma biológico, como por los particulares logros obtenidos en el análisis del pensamiento de algunas de las grandes figuras que contribuyeron a la elaboración de dicho paradigma, por ejemplo, de Henle y Koelliker. Otro gran acierto ha sido, a mi parecer, el carácter didáctico con el que el autor ha pretendido acercarse al problema. Son muy de agradecer por el lector, *verbi gratia*, las recapitulaciones que, al término de la exposición del pensamiento de un autor concreto, incluye como síntesis. Ese mismo espíritu se advierte también en el capítulo *Fuentes básicas para la historia de la Teoría Celular* que se incluye en las páginas finales de la monografía (pp. 289-291).

Sin embargo, debemos señalar algunas ausencias que empañan muy secundariamente el valioso juicio que acabamos de emitir sobre la obra. En primer lugar, echamos en falta un capítulo monográfico dedicado a la evolución histórica de la tecnología en el desarrollo de esta teoría, tanto por lo que se refiere a los cambios que experimentaron el microscopio y otros instrumentos ópticos, como a las mejoras de las técnicas histoquímicas. En segundo término, con objeto de valorar adecuadamente la problemática hispana, hubiera sido de desear un capítulo, también monográfico, dedicado a las aportaciones de los científicos españoles a este paradigma y a la recepción de las teorías foráneas en nuestro país. A pesar de estas ausencias, menores, como ya dijimos, nos encontramos con una notable obra que viene a enriquecer señaladamente la literatura mundial publicada hasta la fecha y que, además, recomendamos vivamente a todos los amantes de la historia de la ciencia. (G. O)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

ARMAS ANTISATELITES

THE SOVIET YEAR IN SPACE: 1982. Nicholas L. Johnson. Teledyne Brown Engineering. Colorado Springs, 1983.

ANTI-SATELLITE WEAPONS: ARMS CONTROL OR ARMS RACE? Union of Concerned Scientists. Junio, 1983.

SOVIET MILITARY PROGRAMS AND THE "NEW HIGH GROUND". Stephen M. Meyer en *Survival*, vol. 25, n.º 5; septiembre-octubre, 1983.

IMPLICATIONS OF ANTI-SATELLITE WEAPONS FOR ABM ISSUES. Donald M. Kerr, Jr., en *Outer Space—Can Militarization Be Stopped?* dirigido por B. Jasani. Stockholm International Peace Research Institute. Taylor & Francis Ltd. En prensa.

METALURGIA SUPERFICIAL PRECOLOMBINA

PRE-COLUMBIAN METALLURGY OF SOUTH AMERICA. Dirigido por Elizabeth P. Benson. Dumbarton Oaks Research Library and Collections, Trustees for Harvard University, 1979

A PRE-COLUMBIAN TECHNIQUE FOR ELECTROCHEMICAL REPLACEMENT PLATING OF GOLD AND SILVER ON COPPER OBJECTS. Heather Lechtman en *Journal of Metals*, vol. 31, n.º 12, páginas 154-160; diciembre, 1979

ANDEAN VALUE SYSTEMS AND THE DEVELOPMENT OF PREHISTORIC METALLURGY. Heather Lechtman en *Technology and Culture*, vol. 25, n.º 1, páginas 1-36; enero, 1984.

EXTINCIONES MASIVAS EN LOS OCEANOS

CRISES IN THE HISTORY OF LIFE. Norman D. Newell en *Scientific American*, vol. 208, n.º 2, págs. 76-92; febrero, 1963.

THE GREAT INFRA-CAMBRIAN ICE AGE. W. Brian Harlan y Martin J. S. Rudwick en *Scientific American*, vol. 211, n.º 2, págs. 28-36; agosto, 1964.

THE WINTERS OF THE WORLD: EARTH UNDER THE ICE AGES. Dirigido por Brian S. John. John Wiley & Sons, Inc., 1979.

GEOLOGICAL IMPLICATIONS OF IMPACTS OF LARGE ASTEROIDS AND COMETS ON THE EARTH. L. T. Silver y P. H. Schultz. Geological Society of America Special Paper 190; 1982.

ELEMENTOS GENETICOS TRANSPONIBLES DEL MAIZ

THE CONTROL OF GENE ACTION IN MAIZE. Barbara McClintock en *Brookhaven Symposia in Biology*, vol. 18, págs. 162-184; 1965.

CONTROLLING ELEMENTS IN MAIZE. Nina V. Fedoroff en *Mobile Genetic Elements*, dirigido por James A. Shapiro. Academic Press, 1983.

ISOLATION OF THE TRANSPOSABLE MAIZE CONTROLLING ELEMENTS *AC* AND *DS*. N. Fedoroff, S. Wessler y S. Shure en *Cell*, vol. 35, n.º 11, págs. 235-242; noviembre, 1983.

PREFERENCIA ATOMICA ENTRE IZQUIERDA Y DERECHA

PARITY NONCONSERVATION IN ATOMS: STATUS OF THEORY AND EXPERIMENT. E. N. Fortson y L. Wilets en *Advances in Atomic and Molecular Physics*, vol. 16, págs. 319-371; 1980.

OBSERVATION OF A PARITY VIOLATION IN CESIUM. M. A. Bouchiat, J. Guena, L. Hunter y L. Pottier en *Physics Letters*, vol. 117B, n.º 5, págs. 358-364; 18 de noviembre de 1982.

ECOLOGIA DEL BOSQUE ESCLEROFILO MEDITERRANEO

BIOGEOCHEMISTRY OF A FORESTED ECOSYSTEM. G. E. Likens, F. H. Bormann, R. S. Pierce, J. S. Eaton y N. M. Johnson. Springer-Verlag; Nueva York, 1977.

EL FLUJO DE ENERGÍA EN UN ECOSISTEMA DE BOSQUE. J. R. Gosz, R. T. Holmes, G. E. Likens y F. H. Bormann en *Investigación y Ciencia*, n.º 20, págs. 47-57; mayo, 1978.

PROBLÈMES D'ÉCOLOGIE: STRUCTURE ET FONCTIONNEMENT DES ÉCOSYSTÈMES TERRESTRES. Dirigido por M. Lamotte y F. Bourliere. Masson; Paris, 1978.

LA CLOROFILA EN LOS ENCINARES DEL MONTSENY. C. Gracia. Tesis doctoral. Universidad de Barcelona, 1983.

CIRCULACIÓN DE NUTRIENTES EN ALGUNOS ECOSISTEMAS FORESTALES DEL MONTSENY (BARCELONA). L. Ferrés, F. Rodà, A. M. C. Verdú y J. Terradas en *Mediterranea* 8. En prensa.

ESTRUCTURA Y EVOLUCION DE LOS GRANOS INTERESTELARES

COSMIC DUST. Dirigido por J. A. M. McDonnell. John Wiley & Sons, Inc., 1978.

DUST IN DENSE CLOUDS: ONE STAGE IN A CYCLE. J. Mayo Greenberg en *Submillimetre Wave Astronomy*, dirigido por J. E. Beckman y J. P. Phillips. Cambridge University Press, 1982.

COMPLEJOS GIGANTES DE NUBES MOLECULARES EN LA GALAXIA. Leo Blitz en *Investigación y Ciencia*, n.º 69, págs. 48-58; junio, 1982.

EFUSIONES DE ENERGÍA EN ESTRELLAS JÓVENES. Charles J. Lada en *Investigación y Ciencia*, n.º 72, págs. 44-55; septiembre, 1982.

ICES IN SPACE. J. Mayo Greenberg, C. E. P. M. van de Bult y L. J. Allamandola in *Journal of Physical Chemistry*, vol. 87, págs. 4243-4260; 1983.

TUERCAS Y PERNOS

THE HERITAGE OF MECHANICAL FASTENERS. Industrial Fasteners Institute, 1974.

STANDARD HANDBOOK OF FASTENING AND JOINING. Dirigido por Robert O. Parmeley. McGraw-Hill Book Company, 1977.

JUEGOS DE ORDENADOR

THE SCIENCE OF SOAP FILMS AND SOAP BUBBLES. Cyril Isenberg. Tieto Ltd.; 5 Elton Road, Clevedon, Avon (England); 1978.

COMPUTERS AND INTRACTABILITY: A GUIDE TO THE THEORY OF NP COMPLETENESS. Michael R. Garey y David S. Johnson, W. H. Freeman and Company, 1979.

TALLER Y LABORATORIO

THE HEAT PIPE. G. Yale Eastman en *Scientific American*, vol. 218, n.º 5, págs. 38-46; mayo, 1968.

EXPERIMENT IN THE BRAGG REFLECTION OF LIGHT FOR THE UNDERGRADUATE USING CHOLESTERIC LIQUID CRYSTALS. A. Olah y J. W. Doane en *American Journal of Physics*, vol. 45, n.º 5, págs. 485-488; mayo, 1977.